

29 May 1980

Watching other people's television

THAT nations, especially new nations, equate their independence with the existence of a national airline is well-known. A separate broadcasting service has more recently also come to seem a symbol of nationhood. It is not merely that heads of state, or their chief executives, consider themselves diminished and demeaned if they cannot from time to time address messages of some kind to their people. Nobody will deny the social and political value of a national television service, potentially at least, in helping to reinforce people's sense that they belong to the same country.

Unfortunately, however, the chauvinism that argues the benefits of national broadcasting services is all too often matched by a xenophobia that amounts to downright suspicion of other people's broadcasts. So much, for example, is apparent in North America, along the forty-ninth parallel — a circumstance that fortunately does not prevent privately-owned Canadian television networks from re-distributing signals picked up from across the border. Along the frontiers within Western Europe, indifference and polite curiosity are probably the hallmarks of people's regard for such neighbouring television transmissions as they can pick up. The frontier between East and West Germany is, on the other hand, much more powerfully charged with television xenophobia. Hitherto, however, sources of strife such as these have been relatively few and far between, largely because of the limited range of ground-based television transmissions. There is now a prospect that this will be changed with the arrival of the first direct-broadcasting satellites some time in the early 1980s.

Already, in Britain, Sir Harold Wilson, the ex-Prime Minister, has talked of the way that British cultural traditions may be undermined by television transmissions (accompanied by corresponding advertising material) from mainland Europe. Elsewhere, similar fears are matched by a sense of opportunity — or even by the hope of commercial profits from novel television channels carrying a mixture of news, entertainment and advertisements to very large numbers of people. Moreover, there is no reason to suppose that the plans which have been laid for the setting-up of direct-broadcasting services are in any way defective.

Geosynchronous satellites operating at centimetre wavelengths will be able to transmit signals powerful enough for television pictures with good definition to be received on individual parabolic dishes one metre or so across. Both in Canada and France, systems such as these are intended to provide cheaper and more effective means of broadcasting existing national services than can be had with the present cumbersome system of ground-based transmitters, land-lines and relay stations. So what are the international implications of these developments? And how valid is Sir Harold Wilson's concern?

First, it must be acknowledged that the system for regulating the use of direct-broadcasting satellites internationally is, bluntly, a muddle. By international consent, the regulatory body is the International Telecommunications Union, which was a great success in the first few decades of its history, when the problems were largely those of arranging that the same countries did not use the same frequencies for long-distance radio transmissions, long or short-wave. Since then, unfortunately, demands on the electromagnetic spectrum have increased, while novel technical developments such as the direct-broadcasting satellite have failed to conform with the precedents for international compromise worked out in the early years.

Faced, two years ago, with conflicting requests for spaces for satellites along the geosynchronous orbit, the technical committees of the ITU for all practical purposes threw up their

hands in despair and chose to allocate to each member state the equivalent of five channels of television broadcasting from geosynchronous orbit. Neither need nor the likelihood that the facilities thus made available would be used was part of the calculation. The result is that in some parts of the world (North America, for example) the would-be broadcasters are unduly constrained, while elsewhere opportunities are likely to be neglected for some time to come. Moreover, while the consents which have been made by the ITU are accompanied by technical definitions of the strength of the signals from broadcasting satellites, so that the footprints on the surface of the earth will be constrained, signals from orbit will inevitably spread further across national frontiers than signals from terrestrial transmitters.

So much is clear. Technically, the first broadcasting satellites could be in service as soon as 1983, both in Europe and North America. It is much less certain what uses will be made of the new technology, and whether the international consequences will be anything like as serious as some have suggested. The fear, for example, that broadcasting satellites might be used to flood various parts of the world with propaganda from another have been to a large extent undermined by the willingness of the member governments of the ITU to accept the technical restrictions put on the footprints from "their" satellites — and by the recognition that a whole population would have to find a way of concealing a metre-dish in every attic before propaganda satellites could be effective.

The more immediate international consequences of the new technology are likely to be in Western Europe, a region divided by its Common Market. Sir Harold Wilson's instincts are to that extent correct. At least in the short-term, however, his fears are misplaced. Britain is divided from the mainland not merely by the English Channel but by the technical specification of its television signals, and from both France and Germany by language. Much the same is true elsewhere in Europe — which is not universally a cause for self-congratulation. It follows that television signals from other people's satellites will be a serious cultural problem only if and when steps are taken to devise not merely their characteristics but their content so that they will compell and hold the attention of Europe's polyglot population.

This will be no light undertaking. One of the attractions of satellite broadcasting is its cheapness. A few tens of millions of dollars will put a satellite in orbit, while the cost of servicing it from the ground with signals is unlikely to be much greater over a period of five years, say. Television broadcasters are, however, these days painfully aware that the cost of burdening the signals their transmitters put out with a content that people want to look at is very much greater. Both in Switzerland and the Low Countries, there are still unformed plans for launching (in both senses of the word) such television services, using channels allotted by the ITU to governments that will, in all probability, be unused. Given that the success of ventures such as these will depend entirely on the skill with which the content of what is broadcast is devised, can anybody make a case that they should be crabbed? Would it not, rather, be more generous and far-sighted to agree that a successful international television service must be inherently desirable and that it should therefore be encouraged? The only basis for holding to a different view must be a conviction (for which there is no evidence) that all television is bad, so that more must be worse.

If the immediate consequences of television satellites are likely to be less serious than people have been suggesting, it does not unfortunately follow that nothing should be done. Nationally, governments will have to look to the arrangements they use at



present for controlling the content of national television broadcasting. Over the years, all liberal governments have laid down rules for guiding the operations of those with a licence to broadcast electromagnetic signals — and from time to time they have prosecuted those who have chosen to violate their airspace as if they were literally pirates. If international broadcasting proves financially feasible, governments will have to reconcile themselves to having less direct control. (They may even be faced with the prospect that national politicians denied the exposure they would like on their national network would go off and hire time elsewhere, a development that might serve to put political broadcasting in perspective.) But this is a prospect that must in any case be faced. Other technical developments, cable television

for example, are similarly a threat to the traditional paternalism of governments and their regulatory bodies.

It will also be necessary to look again at the mechanism by which the ITU sets out to allocate positions along the geosynchronous orbit. For more than the coming decade, and until still higher electromagnetic frequencies are useable, the geosynchronous orbit will be a scarce resource. It is absurd that it should be shared out like a pack of cards. Means of counting need in the calculations should urgently be sought. And some thought should be given to the use of satellites as a means of making television signals internationally accessible by individuals with the wish to enjoy this luxury (and the willingness to pay for it). Television as such is valuable.

Let Finniston cool his heels awhile

THE British Government is taking its time in responding to the report of the Finniston Committee on engineering education published last year (*Nature* 279, 352; 1979), and with good reason. The Department of Industry is planning to make public some kind of an opinion late this summer, but the Department of Education and Science will still then be taking outside opinion. It is exceedingly unlikely that the government will have hit on a policy before this year is out. Others than the impatient members of the Finniston Committee, some of whom appear to have mistaken their report for the Mosaic Tablets, will welcome a breathing space before an important part of British higher education is thrown in the melting-pot.

British governments, like their taxpayers, have been worrying about engineering education ever since the end of the Second World War. At first, the chief concern was that universities were not producing enough graduate engineers to take their places alongside those trained in more traditional ways, essentially by apprenticeship. That defect has now been remedied and the annual output of graduate engineers (from polytechnics as well as universities) approaches 20,000. Concurrently, its traditional route has been closed. More recently, interest has centred on the character and even the quality of engineering education, provoked chiefly by the reflexion that the faults of British manufacturing industry must somehow be attributable to the engineering professors up and down the country.

Few would argue that British engineering education is blameless. It is, for example, difficult to understand how British universities are confident that three years of higher education will enable a man or woman to function as an engineer (albeit under the supervision for a time of one of the engineering institution) when most other educational systems consider three years insufficient. There is also ample if anecdotal evidence to suggest that British engineering education is too much dependent on lectures, too theoretical and too much detached from the practical problems of practising engineers. Not even the Finniston Committee has thrown light on questions such as these, however. Although it is possible to understand why it concentrated instead on the problems of how best to enhance the public prestige and the self-esteem of the engineering profession, a valuable opportunity has thus been missed. And the engineering departments, conscious as many of them are of the need for qualitative change, have no solid foundation on which to base reform.

While the Finniston Committee was sitting the initiative was stolen by the University Grants Committee, which announced two years ago a scheme under which selected universities (seven in total) would be given extra resources in order to provide four-year engineering courses for more able students differing from those of conventional pattern in that their curriculum would be "enriched" with elements of management science, industrial

relations and the law. The UGC courses were also devised so that students would spend some of their time gaining first-hand experience of what was considered two years ago to be the sector of British industry most in need of improvement — manufacturing industry. A few years from now, when the first graduates of these courses are in jobs, it may be possible to tell whether the experiment has been successful, although the chances that an objective assessment will be possible are diminished by the lack of formal criteria by which that might be done.

On the engineering curriculum, the Finniston Committee accepted that the UGC's experiment should be the model for training the most able engineering students. To that extent, the experiment is an experiment no longer. The committee's chief educational concern was to ensure that there would be funds enough to support such courses (and the student following them) wherever in Britain engineering is taught. To this end, the committee asked that its proposed Engineering Board, whose function it would be to supervise professional standards, should also have at its disposal funds with which to supplement existing ways of channelling public money to universities and students.

This proposal is beguiling but is also a trap. Naturally it has the appearance of virtue at a time when the belief persists that a *something*, almost anything, must be done. Understandably, engineering departments are also sympathetic to the notion that more funds should come their way. But changes of the kind proposed, while helping to change the education of engineers in a direction not yet proven, would certainly change the character of British universities in a way that cannot be welcome. In their essentials, these Finniston proposals are tantamount to asking that engineering departments in British universities should be dealt with differently — and more generously, than other departments, and that engineering students should often receive higher stipends. Little imagination is needed to guess how divisive these provisions would be. And there is no evidence that their promised benefits would materialise. It is no wonder that the government's response is hesitant.

Already there are signs that British universities are more jealous of their autonomy than they are skilled at exercising it. Will their interests in the long run be served if an important part of their contribution to the national interest is controlled and financed from outside? (Although there are similarities with medical education, Finniston's proposals for engineering education go much further than the General Medical Council would dare dream.) May it not be that the universities' best course of action, in this as in other matters, would be to carry out their own more constructive examination of the problem? Nobody would be surprised if it then turned out that what is wrong with British engineers is that industry makes too little use of them.

Risk assessment by numbers

Washington

Is comparative risk assessment an idea whose time has come? Or a political attempt to short-circuit necessarily complex sets of rule-making procedures?

The answer seems to be a bit of each. Professional scientific organisations such as the American Chemical Society are now pushing for comparative risk assessment as a way of rationalising government controls on technological activity, and Congress, responding to industry's demands to lighten the burden of controls, is looking at the type of legislative backing needed to increase its role in the regulatory apparatus.

Congressman Jim Ritter of Pennsylvania, for example, has introduced a Bill that would require the Office of Science and Technology Policy to set up a government-wide mechanism for assessing the comparative risks involved in actions in "scientific, technological and related fields".

The Carter Administration is reluctant to enforce centralised risk assessment as a federal policy. But individual agencies are already giving it increasing prominence within their administrative procedures, accepting that the absolute standards of zero-risk, popular with law-makers in the mid-1970s, may be neither practicable nor scientifically sound.

The push for comparative risk assessment is part of a general drive to restructure the regulatory process by separating the scientific judgment of whether a substance or process poses a health hazard from the administrative judgement of how this hazard, once characterised, should be dealt with.

Not all agree that such a separation is desirable. "The location of a centralised entity for performing comparative risk assessment within OSTP undermines the basic consideration that risk management decisions are social policy decisions which are based only partly on hard science", Dr Nicholas Ashford of the Massachusetts Institute of Technology told the House Science and Technology Committee two weeks ago.

Others, however, including the powerful National Association of Manufacturers, see it as an essential step towards rational regulation. "Before a regulator is faced with deciding how much the public needs to be protected from a substance, he or she should be supplied with an objective assessment of the degree of risk which the substance provides", Dr William J McCarville, Director of Environmental Affairs for Monsanto, told the same committee.

Prompted by such concerns, William Wampler, senior Republican member of the House Agriculture Committee, has introduced another Bill to set up a National

Science Council. This would also operate under the auspices of OSTP, and would be responsible for deciding questions of scientific fact arising from debates on the potential harmfulness of substances involved in food production, decisions which would be binding on all federal agencies.

A similar proposal is being canvassed by the American Industrial Health Council, a body of over 100 chemical companies and trade associations, which wants the National Academy of Sciences to set up a panel of eminent scientists to evaluate the qualitative and quantitative aspects of human chronic health risks associated with all forms of industrial activity.

Neither proposal has yet provoked much enthusiasm within the Administration. But on the question of comparative risk assessment, there is a greater consensus that this should play a more prominent role in decision-making — although reluctance to impose it in a too heavy-handed way, stressing instead the need to build up the research capabilities of the regulatory agencies.

OSTP officials point out that several moves have already been taken in this direction. Two years ago, for example, in issuing an executive order on regulatory reform, President Carter directed that analyses of the costs and benefits of proposed regulation should be prepared before the regulation was introduced.

Furthermore, procedures for standardising the scientific evaluation of hazards across the federal government are being developed by the Interagency Regulatory Liaison Group which has already produced a uniform regulatory policy on carcinogens.

But there remains the problem that different agencies have had their legislative mandates written at different times and in different circumstances, giving rise to what OSTP Acting Associate Director Denis Prager describes as a "crucial and necessary" pattern of diversity and pluralism in dealing with technological risks.

The US Department of Agriculture, for example, must observe meat and poultry inspection acts which state unambiguously that no substance, whatever its benefits, may be added to either food if it poses any risk to human health.

Similarly, the Occupational Safety and Health Administration has no legal requirement to take factors other than the safety of workers into account in devising exposure levels to toxic substances (although this could change as the result of an imminent Supreme Court ruling on whether cost-benefit calculations should be taken into account).

Perhaps the greatest push for comparative risk assessment has come in the nuclear field, and in particular over

arguments that base their support for nuclear power on a direct comparison with the health and environmental risks from other energy sources.

At the Environmental Protection Agency, however, which perhaps more than any other agency has had to develop different risk assessment procedures for different types of hazards, there is doubt that comparative risk analysis can be tuned sufficiently finely to meet the needs of specific regulatory situations.

"A comparison of risks within a particular regulatory context may be a useful decision-making tool, and we at EPA are ready to use it as such", Dr Richard Dowd, Chairman of the Agency's Science Advisory board, told the Congressional committee, "but I question how useful it is to go beyond the context of one statute to several statutes, or from one risk situation to a different risk situation."

David Dickson

Research animals

Two bills too many

REGULATIONS to control the use of animals in scientific experiments in Britain are likely to change within the next year or two. Before it took office in April last year, the Conservative Party pledged itself to update the existing legislation, based on the Cruelty to Animals Act of 1876. At the end of last year, Mr Timothy Raison, Minister of State at the Home Office, reaffirmed that intention when he said that the government would like to base the new legislation on the European Convention on laboratory animals, now being drawn up in Strasbourg.

Home Office officials estimate that the European Convention will not be ready until April, 1981. This means that the new UK legislation proposed by the government could not be introduced until the middle of next year at the earliest. Not everyone is prepared to wait that long. Two individual members of parliament — Lord Halsbury and Mr Peter Fry — have prepared their own bills for consideration by the Houses of Lords and Commons respectively. Debates on both bills have been proceeding throughout this session of Parliament.

Lord Halsbury's bill is further than Mr Fry's in the parliamentary process. It received its second reading last October and has just been through several months of scrutiny by a House of Lords select committee. An amended version of the bill, based on the select committee's recommendations, was published last week.

Mr Fry's bill is currently embroiled in debate in a House of Commons standing committee. The committee has so far amended the first six and three quarters of the bill's 38 clauses. At the current rate of amendment the committee will be sitting until August 1982.

Parliamentary procedure, however, will not allow Mr Fry's bill to last that long. It was introduced as a private member's bill and, as such, has to be passed into legislation this session if at all. If it fails to complete the course by late summer, it will have to go back to compete with other private members' bills in the ballot box for the next session.

The Halsbury Bill seems the more likely of the two to survive the parliamentary hurdles. It is also more likely to have the support of the scientific community because it does not advocate prior judgement of the value of scientific research with animals. Its amended form acknowledges and builds on the Protection of Animals Act 1911 but would repeal the Cruelty to Animals Act 1876. The 1911 Act prohibits general cruelty to animals, such as kicking, beating or over-working, and deals specifically with laboratory animals.

In re-drafting Lord Halsbury's Bill, the House of Lords select committee has also produced a report summarising the evidence it has taken and the need for new legislation. The report, which was made available to *Nature* last week, but is not yet publicly available, says that new legislation is needed because of the great increase in the number of animals used in laboratories over the past century and the greater variety of procedures, not in practice in 1876, to which they may be subject.

Lord Halsbury's Bill follows the 1876 Act in requiring all experimenters with animals to obtain a licence from the Home Office. However, it differs in demanding greater specification by those applying for a licence of the number and type of animals to be used and the procedures to be done on them. Applicants for licences would also be required to state the degree of pain or suffering likely to be experienced by the animal and the measures that would be taken for relieving it.

Lord Halsbury's Bill also incorporates the 'pain' and 'anaesthesia' conditions. These are that an animal that appears to be suffering severely under any procedure must be given immediate relief or be humanely killed, and that animals undergoing procedures which may be expected to cause severe pain or suffering must be properly anaesthetised before the procedure starts, be relieved of any pain afterwards or killed humanely. If this is not feasible it should be humanely killed. The 'pain' and 'anaesthesia' conditions are not written into the 1876 Act, though they have been effectively enforced by the Home Office cruelty to animals inspectorate.

Other points of difference between the 1876 Act and the Halsbury bill concern

- the nature of the sponsors. Both require applicants for licences to have two sponsors; the 1876 Act says that one must be a professor of biology, the other a president of a learned biological society; the Halsbury Bill says that the two sponsors must be chosen from a panel nominated by the Secretary of State.

- the term of licences. The Act allows licences to run indefinitely, the Halsbury Bill for only five years.

- the constitution of an advisory committee. The 1876 Act does not provide for an advisory committee on laboratory animals, although one has been in operation since 1906. The Halsbury Bill would incorporate the advisory committee in the legislation and would give it greater powers than at present. For example, the committee would be required to monitor the extent to which animals are used in painful or stressful procedures, review the possibility of alternatives, and inform the Secretary of State on public opinion. It would also have to consider ethical matters.

- the function of a code of practice. One does exist, although it is not written into the 1876 Act. The Halsbury Bill would incorporate it, increase its status to the equivalent of the highway code in Britain. Although infringement of the code would not constitute an offence in itself, it could be used as evidence against someone suspected of malpractice.

- batch testing of sera and vaccines. This is not included under the 1876 Act but would be under the Halsbury Bill. Under this provision, the number of procedures notified to the Home Office and the number of licences would increase considerably. Although the idea behind the Halsbury Bill is to include procedures not covered at present, most survey, field, and school work would be exempt. For example, simple injecting of animals or ringing birds would not be included.

- species of animals covered. The 1876 Act covers all vertebrates; the Halsbury Bill would extend protection to chordates and embryos or larvae capable of independent existence.

The Halsbury Bill is also broader in scope than the existing legislation. But it does not specify detail, preferring to leave that to the secretary of state. For example, he would be able to extend the bill to include other species if he chose. He would also appoint the advisory committee.

The amended Halsbury Bill will go to a committee of the whole House of Lords on 20 June. Once it has been through the report stage, it can be presented to the Commons but it is unlikely that the Commons will consider it before next session. It could then be thrown back at the Lords. Although some believe that new legislation could be passed this year, others believe that the government is likely to stick to its plan of seeing the European Convention first.

Judy Redfearn

X-rays

Selling a new source

THE European Science Foundation — normally a quiet body conducting its catalytic business behind the closed doors of committee rooms — is supporting a proposal that Europe should spend £55 million on a new, 5 GeV super-bright European Synchrotron Radiation Source (ESRS). Last week at Daresbury, UK, it tried to convince the British synchrotron users that it was right. The result was cautious approval.

Britons are probably the hardest to convince of the case. Their own dedicated source, the SRS, will start producing data at Daresbury later this year, and it will be well ahead of other European sources (*see figure*). Moreover their purse-holder, the Science Research Council, is committed to other large accelerator projects (the Nuclear Structure Facility at Daresbury and the Spallation Neutron Source at the Rutherford Laboratory) up to 1984. And the cash available to run the SRS will only be enough to operate half its beam ports.

The challenge, though, comes from the US. The £12 million National Synchrotron Light Source — a two-stage source under construction at Brookhaven Laboratory, Long Island — will come on line in July 1981 (0.7 GeV) and October 1981 (2.5 GeV). The second stage should exceed the brilliance of the 2 GeV, £5 millions SRS by a factor of ten — giving the SRS just a year as world leader.

So Yves Farge, chairman of the ESF *ad hoc* committee on synchrotron radiation and director of the DCI synchrotron radiation laboratory at Orsay, France, argued last week that Europe will need a more advanced source if it is not to be in the second league. Pressure is also high from countries without access to a national synchrotron source — particularly the Scandinavians, Belgium and The Netherlands — for a European facility to be built.

According to Farge, the Scandinavians — will make the construction of the ESRF a condition of their acceptance of LEP, the next high energy physics accelerator contemplated by the European centre for nuclear research (CERN). The Scandinavian view, says Farge, is that the ESRF should be constructed in the tunnel vacated when the CERN intersecting storage rings (ISR) are shut down as a preliminary to LEP.

Farge is unhappy about that approach, however. There are no biologists at CERN, and one of the attractions of synchrotron radiation is that it is a general tool of structural analysis useful to most disciplines. Biologists, however, would feel isolated at CERN. It would be better, he

feels, to choose a different site and set up a private association to which national research councils could contribute, avoiding the need for complex international agreements. (There is a similar foundation at the Institut Laue Langevin at Grenoble.)

There is also a rapidly increasing interest in synchrotron radiation in industry, Farge believes. DCI at Orsay profited from industrial participation by £5,000 in 1979, and the figure will be £20,000 this year. Oil companies appear to be interested in the routine use of the light to seek out significant trace elements in core samples for oil prospecting; the X-ray intensity available could give them a factor of 1,000 in speed.

However the ESRF itself is a long way from a final design. A sub-group of Farge's committee, headed by Dr DJ Thompson of Daresbury, has prepared a preliminary study which defines a 5 GeV, 0.5 Amp, 604 m circumference ring with the brilliance function shown in the figure. Drs B Buras and G V Marr have also defined a shopping list of instrumentation. The machine and building plus half the instruments would cost £55 million at present prices, and could be constructed within 3 to 6 years.

But a new idea has occurred to the committee: that it be an "all-wiggler" machine, a design in which the experimental radiation is taken not from the main bending magnets of the ring, but

from multipole, high field magnets which introduce local high curvature wiggles in the beam path. This would enable the machine to use lower energy and beam current for the same radiation intensity. And the overall radius, as it happens, would be "just right" for the CERN ISR tunnel.

The uses of the ESRF radiation would stretch from molecular biology to nuclear physics, the meeting was told. Using 'undulators' — wigglers with 50-100 poles that produce coherent radiation — spectral brilliances up to seven orders of magnitude greater than those obtainable at DCI are conceivable. Such increases are very rare in science, said Farge, and the consequences are unpredictable.

However Daresbury's Dr Mike Hart, who took the role of the ESRF's chief bubble-burster, claimed one thing is predictable: that such an intensity of radiation from the undulator would destroy most samples. Quick calculations by Farge and Thompson led to estimates of 33-300 Watts onto the sample. "You'll have cooling problems," says Hart.

Probably the most exciting application of the high intensity, coherent light from an undulator would be to illuminate a zone plate X-ray microscope — one that depends on diffraction effects to produce focusing and de-focusing equivalent to the lenses of an ordinary light microscope. This application is limited by the fineness

with which zone plates can be drawn, but object resolutions of 100 Angstroms are foreseeable; and with a wavelength of 25 Å, tuned to detect carbon but not oxygen, the prospect opens of following ultrastructural movements in live cells.

Nevertheless the ESF has not argued the case sufficiently for the ESRF as against existing synchrotron light sources, says Hart; 70% of the arguments used in the ESF document "the scientific case" for the ESRF apply to synchrotron radiation in general, and not to the ESRF in particular. The document will have to be rewritten in two years' time, Farge admits.

Robert Walgate

Four documents on the ESRF are available from the European Science Foundation, 1 quai Lezay-Marnesia, F-67000 Strasbourg, France: 'The Feasibility Study', 'The Scientific Case', 'The Machine', and 'Instrumentation'.

Training

Thoughts from the think tank

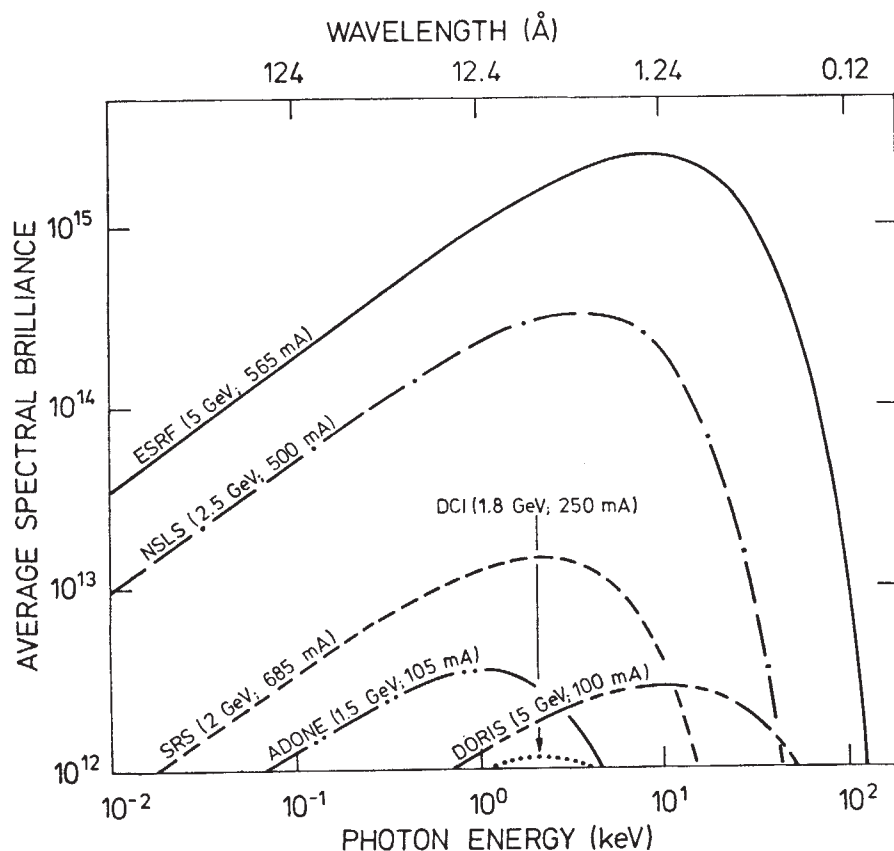
THE British government's own think tank, known as the Central Policy Review Staff (CPRS), published last week a mildly subversive prescription for making the products of the British educational system more suited to what is called "the world of work". The CPRS report ("Education, Training and Industrial Performance", HMSO, £4.25) gives no explanation of how it came to be published or even written.

Among the points at which the recommendations of the report may seem at odds with current government policy, the following are conspicuous:

- There should be formal standards and qualifications for skills acquired by means of vocational training.
- The government should continue to provide vocational training for young unskilled workers.
- More attention should be paid, especially in the age range 16-18, to the acquisition of practical skills, preferably in colleges of further education, not schools.

The CPRS does however accord with current government policy in its advocacy of a core curriculum (consisting of English, mathematics, science, a practical subject and "possibly" a foreign language) in the age range 11-16. The report also asks that the financing of further education by public loans should be reconsidered.

The CPRS is suitably modest in its acknowledgement of how little is known of the relationship between education training of any kind and the eventual benefits to employers and the national economy. There is no single issue on which a government initiative would, in the CPRS's judgment, make training radically more effective — and, even if there were, implementation would be hampered by the



Spectral brilliance (in photons $s^{-1} mm^{-2} mrad$ in 0.1% bandwidth) is compared for the proposed ESRF, the Brookhaven NSLS (ready

October 1981), the Daresbury SRS (October 1980), and the Frascati ADONE, the Hamburg DORIS and the Orsay DCI (in operation).

decentralisation of the British educational system.

Some of the ironies of the British educational system are however neatly described by the report, which is based on an unsystematic survey of the opinions of educational and other organisations and on the CPRS's own views. One large company is quoted as holding the view that university science courses are "too academic" but nevertheless choosing to recruit graduates from the same academic courses in the belief that these would have attracted the most able students.

In its assessment of the provision of education and training, the CPRS echoes the conventional wisdom that Britain is better supplied with means of initial education and training than with facilities for continuing education and retraining.

The report has few concrete suggestions for the improvement of continuing education in Britain. It remarks that elsewhere than in the United Kingdom, and especially in the United States, continuing education flourishes because people acquiring extra qualifications are often rewarded with extra pay, and notes regretfully that "pay policy and union pressure" have prevented such incentives emerging in the United Kingdom. It therefore pins its hope for the future on part-time and evening study.

The more controversial recommendations are, predictably, those that affect the pattern of school education in the United Kingdom. The CPRS comes to the same conclusion as the Secretary of State for Education and Science, Mr Mark Carlisle, that there should be a core curriculum in the age range 11-16 but that this should be adopted voluntarily by educational authorities and not centrally imposed by Act of Parliament. (The report notes that the only present legal compulsion, the requirement that religious knowledge should be taught in schools, "has not been notably successful".)

The CPRS also wants there to be a more deliberate concern by those who design school curricula for the needs of industry

(and is explicitly critical of the School Mathematics Project, now widely used in British schools, for its failure to consult with industry).

On the shortage of science and mathematics teachers in British schools, the report asks that the government should "grasp the nettle" of paying teachers with special skills more than teachers of other subjects. The report does not refer to the longstanding opposition of the teachers' unions to proposals of this kind.

The CPRS also has views about the way in which academics in British higher education should be paid. It refers in its report to the system in the United States under which some academics are paid for only nine months in each year, but are free to take paid employment elsewhere (or to carry out research) in the remaining months.

The CPRS is especially attracted by the way in which this practice may give some academics first-hand experience of industry, and asks that the Department of Education and the University Grants Committee should consider the problem and "use their influence" to get rid of existing restrictions. It hazards no guess at the salary levels that would be established after such a reorganisation. □

Three Mile Island

Reactor entry bid aborted

AN attempt by two Metropolitan Edison engineers to enter the contaminated containment dome of the damaged Three Mile Island reactor was abandoned last Tuesday (20 May) soon after 9.00 am when the inner door of an airlock refused to open. The engineers made three attempts to open the 3' thick steel door leading into the containment dome by the normal procedure of giving the wheel of the door a three quarter turn and pushing. After a

total of 11 minutes of the planned 15 minute "mission", the men were recalled by radio. A total of 10 millicuries of krypton-85 was vented to the atmosphere during the procedure and the engineers received a whole body radiation dose of 10 to 15 mrem.

Officials from Metropolitan Edison in Middletown, Pennsylvania, and General Public Utilities in Parsippany, New Jersey, said that further attempts to enter the dome would be postponed "indefinitely", until the situation could be analysed further. "We have only one theory" said Sandy Solon of Metropolitan Edison. "Corrosion, either around the hinges or around the door seals, is causing the door to stick". The airlock door is made to close tolerance and has not been opened for 14 months.

By noting which instruments inside the dome have not been operating, engineers have estimated that the dome is flooded to a depth of seven feet, "plus or minus one or two per cent". The airlock door is 20 feet above the floor so "there is no danger of a release of radioactive water into the airlock." Video monitors show a steady mist is falling inside the dome as heat from the reactor (which is still at 200°F) causes water on the dome floor to evaporate, rise to the ceiling and condense. "The atmosphere inside is highly corrosive, to say the least" said Bill Murray of GPU.

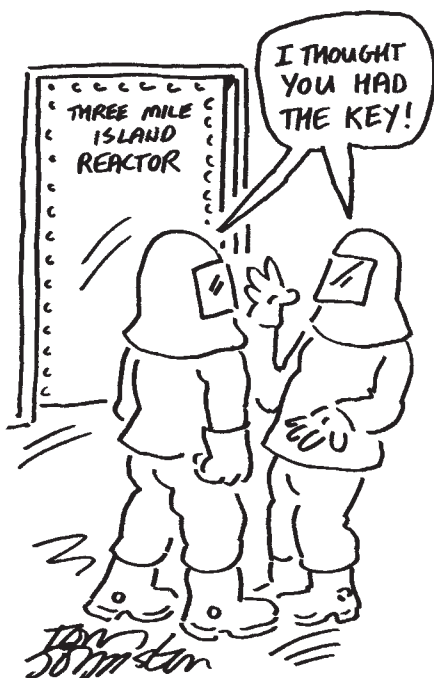
The presence of corrosion in the door "bodes ill" for the functioning of ventilation equipment inside required to cool the dome and maintain a negative pressure with respect to the outside to minimise the possibility of leaks. Engineers have rejected the possibility of forcing the door with hydraulic jacks for fear of tearing the polyplastic inflatable door seals. A failure of the air-lock door to re-seal might necessitate the uncontrolled venting of the 57,000 curies of krypton-85 still trapped in the dome. Bill Murray of GPU said that the Nuclear Regulatory Commission may give the plant operators permission to make a controlled release of the radioactive gas to the atmosphere, thus

Three Mile Island reactor — which way back in?



permitting a forced entry. "Hundreds of thousands of curies have already been released with no ill effect. We hope the NRC will let us vent with maximum controlled dispersion and we can then reassess the situation." Non-gaseous isotopes such as Sr-90 and Cs-137 are presumed to be precipitated on the walls of the dome or trapped chemically in the ventilation filtering system.

The purpose of the attempted 15-minute entry was to monitor radiation in the dome, estimated to be 3 rem per hour of gamma radiation and 200-300 rem per hour of beta radiation. The engineers were also to check for radiation hot spots, examine the ventilation and cooling systems, check for structural damage and take "swipe samples" with gauze of radioactive residues on the walls and stairwells for detailed isotopic analysis.



The two engineers, Michael Benson, aged 27, a nuclear engineer employed at Three Mile Island since 1974, and William H. Berhle, aged 36, a project engineer employed there since 1967, were two of several volunteers. They were fitted with Viking dry diving suits (heavy rubber suits of the type used by helmeted deep sea divers) covered by the standard yellow anti-contaminant radiation protective plastic clothing. Breathing apparatus was a Scott airpack of the type used by US firefighters with a 30-minute air supply. The men practised for two months on the identical airlock of the second Three Mile Island reactor.

Tuesday's failure was the second abortive attempt to enter the building. A previously scheduled entry on 24 April had to be cancelled when it was discovered that modifications in the breathing apparatus prepared by the Bio Marine company had not been submitted to the National Institute of Occupational Safety and

Health for approval. At a press conference on 21 May, Benson and Berhle said they were "disappointed" and would volunteer again. The GPU said it would announce its next move "within a week".

Joe Schwartz

Recombinant DNA

Guidelines for scale-up

Washington

Like a man with a new pair of shoes, the US National Institutes of Health are experiencing some discomfort in administering industry's voluntary compliance with the safety guidelines covering recombinant DNA research.

Pressures on the NIH are coming from two directions. On the one hand, increasing competition from European and Japanese industry is leading US companies to complain that lengthy procedures for approving new host/vector systems and large-scale fermentation experiments are becoming a commercial handicap.

On the other, the members of the NIH's Recombinant DNA Advisory Commission are feeling their way uncertainly into the engineering and safety aspects of large-scale fermentation techniques of which few possess either detailed knowledge or expertise.

The system of voluntary compliance was proposed last year by NIH Director Dr Donald Fredrickson, largely as a way of avoiding the need for new legislation extending the guidelines to the private sector. Since the system was introduced in January, eleven companies have registered their Institutional Biohazard Committees (IBCS) with the NIH, and seven large-scale experiments have been approved.

Last year the containment guidelines were reduced for most experiments to a level requiring minimal physical safeguards, with the result that few companies now feel them to be excessively restrictive. Indeed some take comfort from the fact that Japan has recently introduced the NIH guidelines in their original, more stringent, form.

There is greater concern, however, about the time taken up by NIH's certification procedures, in particular that required to review new host/vector systems — a responsibility which many RAC members now consider to be one of the committee's most important functions.

At present, for example, the disabled bacterium strain *Escherichia coli* K-12 is the only approved host/vector system for experiments at the P1 minimum containment level (although the use of *Saccharomyces cerevisiae* strain of yeast is

expected to be recommended for approval at RAC's next meeting).

The result, according to Dr Peter Farley, President of the Berkeley-based Cetus Corporation, is that European companies who can locate their operations in countries with minimal certification requirements already have a major commercial advantage through being able to exploit other host/vector systems, such as *Bacillus subtilis*.

Time is also being taken up by the need for special approval by the Director of NIH for each experiment carried out using more than 10 litres of culture — a relatively arbitrary limit set in the early days of the guidelines and apparently based on conventional laboratory practices.

Amendments to the guidelines which would accelerate the approval process for commercial developments have now been recommended to NIH by Dr Irving Johnson, Vice-President of Research for Eli Lilly and Company, which has already received permission for several scaled-up experiments in the production of human insulin.

In particular, Dr Johnson is proposing that, once approval has been given for a particular experiment to be carried out above the ten-litre limit, further volume increases using the same biological materials at the same containment levels would require merely to be approved by the local IBC.

More controversial is Dr Johnson's suggestion that, in order to make up for RAC members' lack of experience with large-scale applications, industry representatives should be added to the committee with expertise in areas such as fermentation technology and engineering.

And he is also proposing that a permanent subcommittee be set up to advise the Director of NIH on actions related to large-scale applications, with authority to recommend approval of preliminary plans for large-scale operational facilities.

The various proposals will be discussed by the RAC when it meets next week. Their reception is uncertain.

Committee members are aware that they lack expertise in either worker safety or fermentation engineering. "Many have not been entirely happy with their role" says committee member Dr David Baltimore of the Massachusetts Institute of Technology.

Yet there is disagreement over whether this lack should be made up primarily by expanding the scope of the committee, through increased use of consultants or the addition of new members. Or, alternatively, whether greater involvement should be encouraged from agencies which already have statutory responsibility for such matters, in particular the Department of Labor's Occupational Safety and Health Administration.

Either way, it seems increasingly unlikely that there will be new legislation imposing

the guidelines on the private sector. A bill seeking to do so, and requiring the registration of all industrial activities using recombinant DNA techniques, has been introduced by Senator Adla Stevenson Jr; but it has so far failed to gain any significant support, either inside Congress or without.

David Dickson

Comecon Science

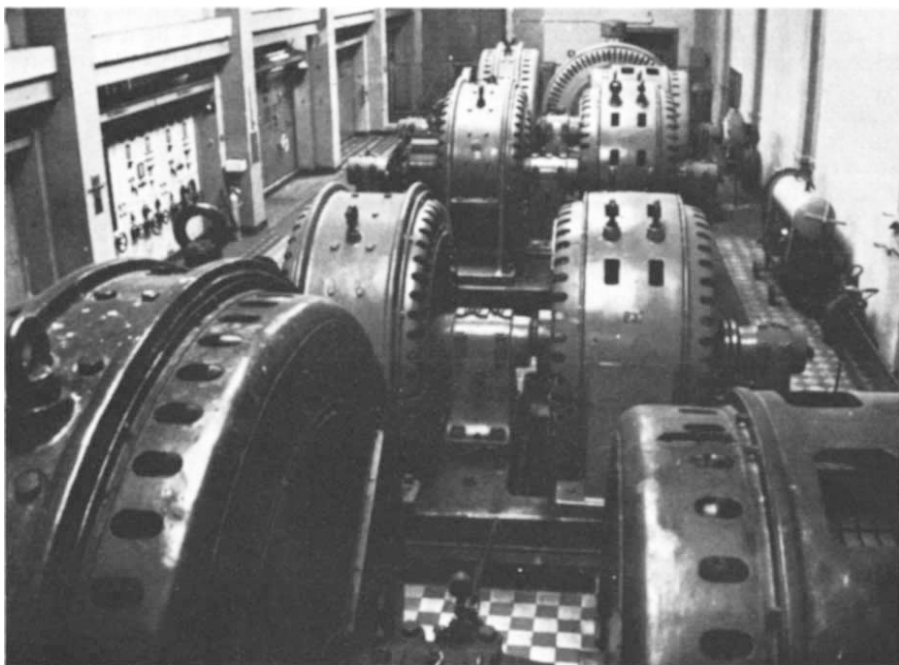
High T, Low K

A NEW Comecon International Laboratory of High Magnetic Fields and Low Temperatures is to be built in Wrocław during the next Five Year Plan (1971-1985) by the Academies of Science of Poland, Bulgaria, the Soviet Union and East Germany, with possible participation also by Czechoslovakia. This will replace the existing somewhat cramped premises which have served the laboratory since 1968.

According to Dr Włodzimierz Trzebiatowski, until recently the Director of the Laboratory, and who still, in retirement, takes a keen personal interest in it, Wrocław was originally chosen as the site of the laboratory for a number of practical reasons. First the Polish Academy of Sciences already had a flourishing Institute of Low Temperature and Structure Research in Wrocław. Wrocław is near the important East German science centres of Dresden and Berlin, and helium, essential for low temperature work, is available from a Polish extraction plant only 100 km away. Finally, just as the laboratory was being planned, the Lower Silesian electricity board, was installing new generators in the Wrocław tramway substation and was willing to donate the old generators to the Academy of Science.

The present laboratory is concerned with such things as the properties of hard semiconductors, new magnetic materials and the measurement of specific heats below 1°K. Cooling below 4.2°K is effected by helium cooling in the field of superconducting magnets of 5T and 15T, and work is in progress on adiabatic demagnetization of nuclear and electron paramagnetics, allowing temperatures below 0.1 K to be attained.

The laboratory has a special interest in compounds of uranium with non-metals of the Vb and Vlb groups — a speciality begun by Professor Trzebiatowski at the Low Temperature Institute of the Academy of Sciences in the early 1950s. (Since his initial discovery of ferromagnetic transition in uranium hydride UH_3 , some fifty such uranium compounds, mostly chalcogenides, have been discovered there.) One new project, of especial interest to Dr Evgeni Leyarovski, of the Bulgarian Academy of Sciences, who was Deputy Director of the Wrocław Laboratory from 1974-1977, is the study of superconductors



Tramway generators keep turning in Wrocław

with magnetic impurities. Research on this phenomenon is also going forward in Sofia, but even with their new 5 tesla superconducting magnets installed this spring, the Bulgarians will still have to do much of their work in Wrocław.

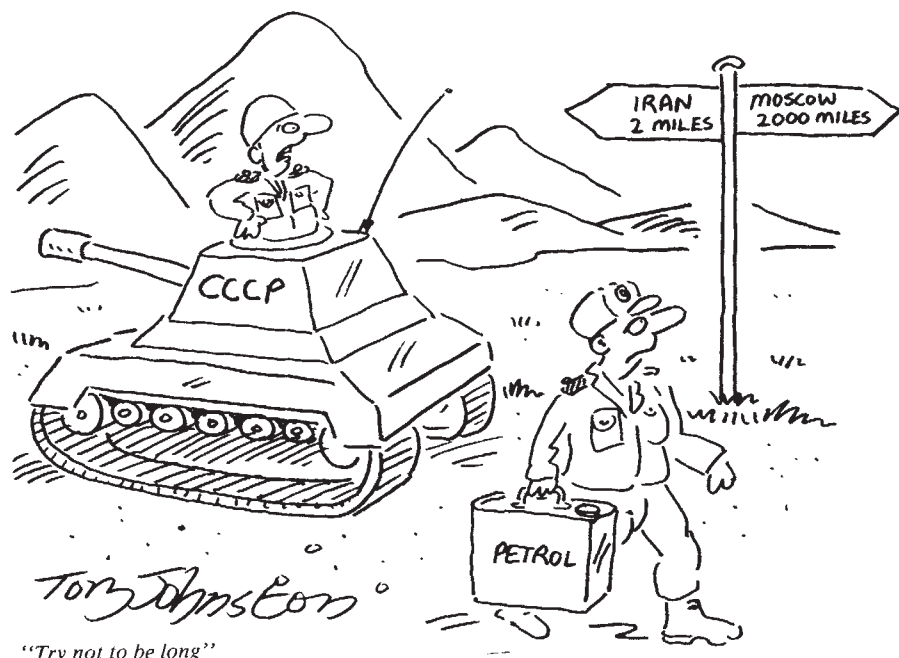
Nevertheless, even Wrocław cannot at present offer the facilities which an international Comecon laboratory needs. Far more laboratory space is required, and, more important, the ex-tramway generators can operate only one of the three magnets at a time, so that the various teams often have to queue for "magnetic time". The new site, on the banks of the Odra, will provide space and facilities for more magnets with higher fields as well as an ample supply of water for cooling the electromagnets.

Vera Rich

Soviet oil

No real shortage

THE Soviet Union has initiated a major change in its method of estimating oil production costs which will result in a dramatic increase in production from proven fields, says a report* published last week by PetroStudies, a Swedish group specialising in the analysis of the Soviet oil and gas industry. The report disagrees sharply with CIA forecasts that the Soviet Union will suffer a shortfall in oil production. The CIA prediction has been used repeatedly in recent months by US oil officials to suggest that a Soviet oil shortage was leading it to have a material



interest in controlling the Middle East.

The PetroStudies group finds that a recent 11-20 fold rise in the reference price of oil, from \$0.75 — \$1.50 to \$17 — \$20 per barrel, will correct the "systematic underexploitation of Soviet oilfields over the past thirty years". The reference price is the figure planners use to assess the cost of opening new oil fields, compared with the cost of intensively developing existing ones. The higher reference price represents a recognition by Soviet planners that development of new oil fields has created unacceptable pressures on capital, labour and material, necessitating a new economic approach to oil development. The PetroStudies report says that the new

policy means that "there is no danger whatsoever that the USSR will be forced to become a net importer of oil this decade and highly improbable that it will become so in the 1990s".

An additional feature of the new policy will be a decline in Soviet imports of American oilfield machinery, the report says. Intensive mining of existing fields by many wells will substitute for high capacity pumping from a few wells, a policy which will significantly decrease Soviet dependence on US-produced high capacity oil pumps.

According to the *International Herald Tribune*, diplomatic sources have suggested that the PetroStudies' findings

may be "part of a Soviet campaign of disinformation". M Jermol of PetroStudies explained that their analysis is based on original Soviet sources obtained through normal institutional channels. The materials include Russian texts of official oil industry documents, specialised books and journal articles, reports from Soviet research institutes, and petroleum conferences, extending back over ten years. "It would be impossible to base a report like this on a few privileged documents" said Jermol.

Joe Schwarz

*"Soviet oil production reform of 1980 and its potential", PetroStudies Co., Sjoblads vag 27, S-21370 Malmö, Sweden. 260pp.

Instruments

New homes for old instruments

Is Britain's scientific heritage being ignored and forgotten by its museums? Arthur Frank, a collector of scientific instruments, believes that it is. About half of his collection of instruments dating from the early eighteenth century, which runs into thousands, has been placed in British museums — mainly central museums in London and Edinburgh. But the other half is on display only in his garage in Jersey, while Mr Frank searches for more conventional homes, so far with little success.

There are more than 1,100 museums in Great Britain and Ireland, but Mr Frank says that fewer than a dozen of them take scientific instruments seriously. But they do appear to differ in their reactions. Some have bought or borrowed instruments from Mr Frank. Others have declined to do so either on the grounds that the instruments which he has available would not fit in well with existing collections or because they would not be able to afford the prices asked. But many, mainly provincial museums, have said that they are the wrong organisation to approach because they do not have the expertise to judge scientific instruments.

One of the few provincial museums to house part of the Frank collection, however, is the Museum and Art Gallery, Doncaster, Yorkshire. Mr J Barwick, the museum's director, had to draw on outside expertise to mount the exhibition. His interest in the Frank collection stemmed from his wish to display a selection of early mining and surveying instruments appropriate to an old mining and engineering town; the public response has now persuaded him also to borrow some of Mr Frank's microscopes and astronomical instruments and he is planning to extend this section of the museum.

Doncaster Museum's special interests may be exceptional. Elsewhere, according to one official, regional and municipal museums are likely to be more interested in a display of instruments representative of different types than of a specialist sub-collection from Mr Frank's garage.

Mr Frank, nevertheless, is determined that his instruments should not be disposed of one by one. He also requires that recipients of objects from his collection should catalogue the instruments and put them all on display. He says that he is prepared to make gifts of them, lend them for twenty-one years, or sell them at two-thirds of the estimated value, dealing with learned societies as well as museums. One snag that may deter some museums is that Mr Frank's valuations seem high.

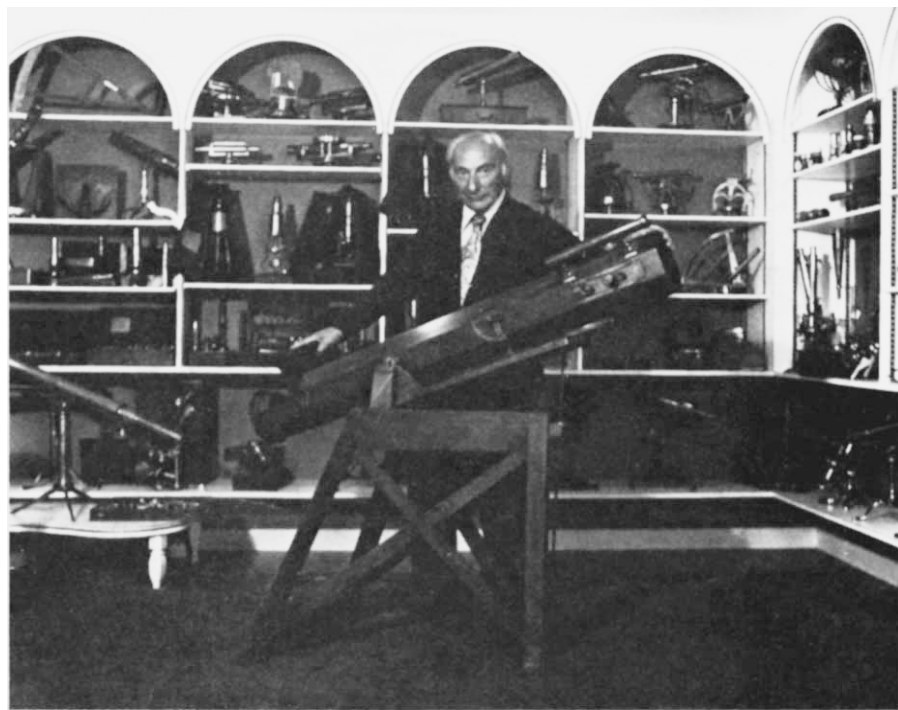
So far, the large specialist and national museums have been the chief recipients either loaning the instruments or more usually buying them under the two thirds

of value offer. The Science Museum in London has about 1,500 of his cameras and accessories, and there is a permanent exhibition of field and opera glasses and of prismatic binoculars. Inevitably, however, such a museum is primarily concerned to fill gaps in its present collection.

Elsewhere in Britain, the Royal Scottish Museum in Edinburgh has a display showing the development of the achromatic microscope in the formative years 1800-1860 (and is also negotiating for the remaining Scottish instruments in the Frank collection). There is also a collection of astronomical and terrestrial telescopes at the Royal Observatory, Edinburgh.

The parts of the Frank collection not yet disposed of are strong on spectroscopic and stereoscopic devices, as well as early spectacles. He is especially proud of his Marshall microscope (circa 1700) and his Ramsden astronomical instruments from the late 1700s.

Judy Redfearn



Arthur Frank with his instruments: gift horse or salesman?

CORRESPONDENCE

Parkinsonian exasperation

SIR, — Further to Robert Moss' letter (1 May, p9,) it is of interest to consider the size of a primitive group of workers when liaison begins to become less effective and an administrator is appointed. Following Parkinson's work on committees (C N Parkinson, *Parkinson's Law*, 1958), experimental evidence from both scientific and socio-political groups seems to indicate that communication remains viable for groups of between 5 to 11, with an optimum of 7 to 9. Between 12 and 14 members, a part-time administrator is usually appointed, and the stage is set for the entry of centralisation.

One of the finest examples of misgrouping to be seen in public was the three day meeting between management and union staff, shown recently on BBC 1. Possibly unaware of Parkinson's work, the two scientists who convened the meeting set up a group of approximately 21 to 23 people, with predictable results. Despite the efforts on both sides to overcome the traditional divisions between management and unions, the lack of communication due to the group size increased exasperation to the point when people began to leave. Amazed by the rising irritation, the scientists sought to alleviate it by altering the seating arrangements (*vide* Parkinson on the shape and size of committee tables). Communication was restored only when the participants spontaneously divided into two groups of 10 to 12 each, both containing managers and unionists. It was difficult to determine who generated this split, given the highly condensed version of events presented, but it was certainly not the scientists. Additional information to confirm or deny the hypothesis that smaller groups were suggested by the women trade unionists would be welcome.

Stage 3 of Parkinson's progression is considered by Moss to produce an 'intolerable' workload on the administrator, when he is burdened by 17 lines of communication

($L_{wi} + L_{ai}$). The above empirical analysis suggests that the maximum tolerable number of $L_{wi} + L_{ai}$ will be of the order of 10 to 12.

Yours faithfully,

B M GRAY

*Climatic Research Unit,
University of East Anglia,
Norwich, UK.*

Feminist speculation

SIR, — The thought-provoking discussion of the reasons why increased centralization is accompanied by decreased effectiveness presented by Robert Moss was both perceptive and pertinent. I should like to comment on an aspect which Dr. Moss appears to have overlooked. He states that the results of one study showed that "the production of scientific papers per man decreased as the organization increased in size." Was the production of scientific papers by women not affected? If not, what are the most likely reasons? Can one assume that women are less likely than men to be affected by this extension of Parkinson's Law?

The administrators described by Dr. Moss were apparently male, and it would be interesting to speculate on the outcome had some or all of the administrators been female.

Yours faithfully,

C RIGBY

*Aylmer,
Quebec, Canada.*

Erratum

In Robert Moss's letter, "Expanding on Parkinson's Law" (1 May) there was a typographical error. In stage four of his administrators' model, the equation $L_{wi} = W$ erroneously appeared as $L_{wi} = W^2 + A$.

Editor, Nature.

Smallpox and conservation

SIR, — It has been announced by the World Health Organization that smallpox now has been eliminated. The virus responsible for smallpox, an organism formerly abundant and widespread, has been systematically, deliberately, and (as it appears) completely eradicated from all the natural systems of the world. With the exception of samples in a few laboratory 'zoos', it is extinct.

Those responsible for the demise of smallpox will be praised universally, I am sure. But what an eloquent dilemma this presents to conservationists.

Statements of the conservation ethic normally include such clauses as "the preservation of species diversity" and "the protection of endangered species". What happens to those ideals in the aftermath of smallpox? Does the conservation ethic inevitably reduce to a pragmatic formula for the protection of the human species, or to a sentimental affection for the cute and the cuddly? Would anyone care to provide a satisfactory statement of what we really do mean when we speak with solemn and genuine concern about the protection of the natural world?

Yours faithfully,

JOHN MIDDLETON,

*Carleton University,
Ottawa, Canada.*

Making books cheaper

SIR, — Your editorial, "Are Books Too Expensive?" (24 April) is a reasonable and most welcome voice in the discussion of rising book prices. At The University of Chicago Press we have been giving a good deal of thought to measures that would decrease the costs of book production and therefore reduce the list price. These include: the publication of more original paperback books; printing some books from camera-ready typescript rather than setting all books in type; asking authors to forego some portion or royalty on the sale of a specified number of copies of books with smaller audiences; and simply not publishing certain classes of books such as the proceedings of symposia because their relevance is commonly limited over time.

The cooperation of those who write, distribute, review, and purchase books is essential to the success of any of these measures. Many scholars prefer to purchase paper-bound books or books that are printed from typescript because they are less expensive. Yet these same scholars balk at the prospect that their own books might be published in such a fashion, chiefly because their colleagues see such methods as characteristic of "second-rate" work or treatment. Some scholarly journals still refuse to review original paperbacks, and some libraries still refuse to buy them.

We as publishers must recognize that what is in the book must reach the intended audience and that the costs of "traditional" production may force the price of a book beyond the means of at least a fair proportion of the audience, particularly those who are students or younger faculty members, but we need and appreciate the cooperation of both authors and readers.

Yours faithfully,

SUSAN E. ABRAMS

*University of Chicago Press
Chicago, UK, US.*



NEWS AND VIEWS

Evolutionary genetics of snails

from J. S. Jones

MOLLUSCS are very suitable for research into evolutionary genetics. They are sluggish (and hence easy to capture and to mark), are common in a variety of habitats, and often display a striking shell polymorphism.

Many evolutionary mechanisms are now known to affect shell polymorphism in the European snail *Cepaea nemoralis*. Their importance varies greatly from population to population and the realisation that the factors influencing even this simple system are very complex is perhaps the most important point gained from the fifty years for which it has been studied.

At a recent conference* M. Lamotte (Ecole Normale Supérieure, Paris) discussed the controversy between himself and A. J. Cain and the late P. M. Sheppard (University of Liverpool); this has often been misrepresented as a conflict between mutually exclusive mechanisms of visual selection by predators, of climatic selection and (in some populations) of genetic drift. However, it is now clear that to explain gene frequencies in individual populations it is usually necessary to produce complex explanations involving several evolutionary forces (for a recent review see *Ann. Rev. Ecol. Syst.* 8, 109; 1977). This complexity probably applies also to the evolutionary biology of other polymorphisms (such as those at loci controlling soluble enzymes) but, as less information is available for most of these, there is still a tendency to search for single, unifying explanations of their existence. Climatic selection is generally accepted to affect gene frequencies in *Cepaea*; S. M. Tilling (North London Polytechnic) showed that different natural and artificial (black or white painted) morphs have behavioural differences which are related to their energy uptake in sunshine. *C. nemoralis* populations living in complex environments have more shell polymorphism than do those from habitats in which most individuals experience the same climatic stress, and experiments with snails marked with spots of dye which fade at a known rate when exposed to the sun suggest that populations exposed to a variety of climatic stresses are indeed more variable than are those with a narrower niche (J. S. Jones & R. F. Probert, University College, London).

Natural selection is also involved in the leucine amino-peptidase polymorphism of the mussel *Mytilus edulis* according to R.K. Koehn (State University of New York). In Long Island Sound a steep cline in allele frequency coincides with a salinity gradient. Each year the cline is displaced to a varying extent by the immigration of oceanic larvae, but each year selective mortality ensures that it returns to the same point.

Historical effects — whose role in the evolution of molecular polymorphism is notoriously difficult to assess — also influence the present-day distribution of shell variation in some *Cepaea* populations. Information from old records on the land use patterns of downland in southern England suggests that some local anomalies in gene frequency result from the spread of populations from woods into open habitats which become available to *Cepaea* only when sheep grazing declined (R.A.D. Cameron, Birmingham University; M. A. Carter, Portsmouth Polytechnic). Changes in gene frequency over the past 20 years in a *C. hortensis* population on the earthwork of Silbury Hill are also correlated with vegetational changes, in this case resulting from increased human usage. The intensity of selection is as much as 10% per generation (S. Wall & M.A. Carter, Portsmouth Polytechnic; B. C. Clarke, Nottingham University). These rapid evolutionary changes are a microcosm of the long-term changes in gene frequencies in English *Cepaea* shown from shells found in archaeological sites by A.J. Cain.

All theories of the origin of species depend on historical events which are by their nature not directly testable. M. J. D. White has recently put forward (*Modes of Speciation* W. H. Freeman, San Francisco; 1978) a theory of 'area effect speciation'. This arises directly from the observation by Cain and Currey (*Phil. Trans. Roy. Soc. Lond. Ser. B.* 246; 1; 1963) of 'area effects' in *Cepaea*. In many places large regions of constancy of allele frequency for shell polymorphism are separated by steep clines (which usually do not coincide with an obvious environmental discontinuity) from adjacent areas in which the *Cepaea* populations have very different allele frequencies. White suggests that each area effect is a marker for a reorganisation of the whole genome and that a slight reinforcement of this could lead to speciation within a continuous population,

without the complete geographic isolation necessary in Darwin's model. This theory adds new interest to studies of genetic differentiation in snails. One prediction might be that if adjacent area effects have diverged at many loci, some of these might be detected by enzyme electrophoresis (as is the case in incipient species of *Drosophila*). R. K. Selander and colleagues (University of Rochester, NY) find no evidence for parallel genetic differentiation at the morphological and the molecular level in some striking area effects in *C. nemoralis* in the Pyrenees. The pattern of distribution of species of the snail *Partula* on the Pacific island of Moorea does suggest that they have arisen as a result of the divergence of adjacent populations. B. C. Clarke and J. J. Murray showed that at some species' borders there are localised hybrid zones between populations which elsewhere act as good species, whilst in one locality there is a 'ring-species' complex with a diameter of only 3 km. Although area effect speciation might indeed be involved here, different species of *Partula* — even those from islands thousands of kilometres apart — show only slight differences in the frequencies of genes controlling molecular polymorphism. The same is true of morphologically distinct species in the freshwater genus *Elliptio* (G. M. Davis, Academy of Natural Sciences, Philadelphia). *C. nemoralis* populations do show genetic structure at the molecular level but this is independent of the area effects themselves. Pyrenean populations assort into several geographically and genetically distinct groups on a factorial analysis of enzyme and shell polymorphism, and there may also be pericentric inversion differences (which could act as isolating mechanisms) between populations of this species (C. Page, North London Polytechnic). However, area effect speciation could scarcely have occurred in the genus *Cepaea*, which has only four species but thousands of area effects.

The lack of concordance of polymorphism in the genes controlling enzyme structure with other indicators of species divergence suggests that other loci — such as regulator genes — might be involved in speciation (Wilson, Maxson & Sarich *Proc.*

*An international conference on Molluscan Genetics was held by University College London and the Linnean Society of London in London from April 17-19, 1980.

J. S. Jones is in the Department of Genetics and Biometry, University College, London.

natn. Acad. Sci. U.S.A. 71, 2843; 1974). The importance of ontogenetic and regulatory effects in molluscan population genetics is emphasised by the work of J. Heller (Hebrew University of Jerusalem) on the Israeli snail *Xeropicta*, an annual species. Adult *Xeropicta* in the mountains are dark coloured, those in the coastal plain pale. As the mean annual temperature is higher in the plain than in the mountains, this accords with the common poikilotherm pattern in which populations from hot places have a higher reflectivity of solar radiation than do those from cool. However, there are great annual fluctuations in temperature in Israel, and in the summer the snails are under extreme solar stress. They adjust their shell reflectivity to ensure that they absorb the requisite energy from the sun during the brief growing season, but are protected from excessive heat at other times; in the mountains they are born pale and become darker, while in the plain they are born pale but become paler. The structural genes controlling shell polymorphism are expressed only at the correct period of the life cycle. Ontogenetic effects are also found in the Caribbean snail *Cerion* (S. J. Gould, Harvard; D. Woodruff, University of California, San Diego). This genus shows a great variety of morphological types. Simulation of the growth process shows that simple alterations of the growth rate early in development can lead to startling differences in adult shell form. Developmental plasticity (perhaps promoted by polymorphism in the genes regulating growth rate) allows each species to attain a multiplicity of adult forms attuned to particular habitats. This may allow a single species of *Cerion* to fill the ecological niches occupied by several species of mollusc on dunes in southern Europe.

Molluscs show a range of sexual strategies from complete outcrossing to obligate parthenogenesis. They are therefore useful in research on the evolution of sex, particularly as the degree of outcrossing can be estimated from the zygotic proportions at loci controlling protein polymorphism. G. McCracken (University of Tennessee) and R. K. Selander discussed the European slugs which have colonised the US. There is great variation in the degree of outcrossing, with some species adopting a mixed strategy of sexual and asexual reproduction. Measurement of the niche breadth of several species on an ecologically well-characterised experimental plot shows that the three species with the broadest niche are present as monogenic and asexual strains. This is hard to reconcile with models which predict that sexual species have more evolutionary flexibility than do asexual.

The conference demonstrated that although the single helix of the snail's shell may be less central to modern genetics than the double helix, it continues to provide important tests of evolutionary theories. □



100 years ago

A proposal has been set on foot for lighting the Sheldonian Theatre, Oxford, and the Camera of the Radcliffe Library with the electric light. It has long been regretted by many members of the University that the Sheldonian Theatre is not available in the evening for any purposes of public interest, however great, for want of lighting. The neighbourhood of the Bodleian Library has, however, been a bar to any proposal for lighting by means of gas or any ordinary method. The care with which the heating apparatus of the Theatre has been inclosed within a fire-proof chamber is sufficient evidence of the importance attached by the curators of the Theatre to absolute security in this respect. The development of the electric light has now rendered it possible to illuminate public rooms by a process absolutely free from danger of fire. It has been adopted largely in the reading-rooms of our public libraries, and notably in the reading-room of the British Museum.

Herr Zehfuss has lately given (*Wied. Ann.*, 4) some personal experiences of the phenomenon of "after images of motion". These after images may be had, e.g., in a train, if one look at a point on the horizon for a little, then turn to look at (say) a horizontal fibre in the wood of the carriage, or close one's eyes. Motions then seem to be still perceived; in the latter case, e.g., a stream of sparks seems to be moving to the right (or if the point originally looked at have been between the observer and the horizon, there is a stream of sparks above going to the right and one below to the left). Herr Zehfuss offers a physiological explanation, in preference to the partly psychical ones proposed by Plateau and Oppel. Each individual nerve rod, he supposes, has special blood-vessels, which, when the original image of a moved object goes to the right, directs the course of the blood to that side, just as in ordinary light the decomposed blood is promptly replaced by fresh. By this preponderant direction of blood to the right a heaping up occurs in each retinal element on the right, which gives rise to return currents as soon as the outer cause has ceased to act. As the blood flows back there arise, in consequence of the specific excitability of the rods, those spark-streams, which are projected as elementary motions to the right.

from *Nature* 22, 27 May, 87 & 90, 1880.

High-resolution imaging of laser-driven implosions

from Lynn R. Veaser

HIGH-RESOLUTION imaging of the radiation (mainly electrons, neutrons, X-rays, and alpha-particles) emitted by laser-driven DT fusion implosions is difficult because such radiation cannot easily be focused. Nevertheless, imaging could be an extremely useful diagnostic tool, especially for comparison with the extensive computer modelling now being carried out for laser-fusion compressions. Neutron imaging has not been achieved because the fluxes are small and the efficiency for detecting neutrons is low. Electrons are generally affected by the strong electric and magnetic fields in the laser plasma but X-rays and, if the fields are not too strong, alpha particles, can be imaged by pinhole cameras.

Unfortunately, to produce a high-resolution image, a pinhole camera must have a tiny aperture and this will cause low sensitivity. If additional signal strength is needed it is necessary to use more than one aperture. The resulting image, or shadowgraph, is a superposition of the images from all the apertures, and it is not useful or recognizable unless a real image of the source, similar to that seen for a single aperture, can be reconstructed from it. The two-step process by which this imaging and the subsequent reconstruction are done is known as coded-aperture imaging.

Coded-aperture imaging has been used by astronomers for a number of years to

image X-ray stars, and it is employed in the field of nuclear medicine to enhance the signal-to-noise ratio and the resolution in radiographing body organs; but, because of the difficulties involved in fabricating the complicated microscopic apertures needed for ultra-high resolution, its application to laser plasma has been delayed for some time. Recently, however, groups from the US laser fusion laboratories at Livermore and Los Alamos have reported the use of coded-aperture techniques in imaging of imploding laser-driven fusion targets.

Ceglio and Larsen (*Phys. Rev. Lett.* 44, 579; 1980) used a Fresnel zone plate (in this case a series of concentric gold rings with radially decreasing apertures and held together by thin struts) as a coded aperture to measure the X-rays from a thin glass microballoon which had been filled with DT fuel and irradiated with about 20 TW of laser power. They obtained X-ray images from the shadowgraphs by optical reconstruction; each shadowgraph was illuminated with a coherent light source, much as is done in holography, producing a magnified image of the initial source. The experiments were done at the Shiva laser facility of the Lawrence Livermore Laboratory.

When a high-power laser pulse strikes a target surface, it generates a plasma which includes a significant number of high-energy (suprathermal) electrons. The

target, consisting of fuel contained in a thin-shell microballoon known as an exploding pusher, implodes when these hot electrons heat the shell, causing it to explode both inward and outward; the inward momentum compressing the fuel. In the Shiva experiment two simultaneous short (typically < 100 ps) laser pulses, each of about equal energy, imploded the shell from opposite points (the poles). X-rays from the implosion were imaged in three energy intervals by stacking layers of X-ray film and filters. Two of the spectra were of great interest. Thermal X-rays (from 4 to 7 keV) were found near the inner part of the imploding shell and are thought to represent the hottest part of the pusher material at a time when its diameter was nearly at its minimum. Two disk-shaped regions, about $40\text{ }\mu\text{m}$ apart and clearly visible in the reconstructed radiograph, showed the pusher temperature to be hottest at the poles. Relatively few X-rays originated at the target center in the region of the compressed DT fuel. Of perhaps even greater interest was an image of the suprathermal X-rays, or bremsstrahlung, derived from interactions of the hot electrons with ions from the exploding pusher. If the hot electron distribution is assumed to be uniform, i.e., if the electron ranges are long compared to most spatial nonuniformities in their production, then a measurement of the bremsstrahlung distribution should map out the ion distribution during the early stages of the implosion. The suprathermal X-ray image (from 17 to 30 keV) published by Ceglio and Larsen shows a hollow shell of nearly the same outer diameter, about $300\text{ }\mu\text{m}$, as the initial shell, but the X-ray density was much more intense near the poles than the equator.

The general shapes of these X-ray images, including the decrease in intensity near the equator, were matched nicely by numerical simulations. Having experimental pictures to compare with the simulations has made it possible to test the importance of various proposed energy absorption and transport mechanisms in the target. The quality of the agreement provides considerable satisfaction with the present state of numerical laser-fusion calculations.

Two features of the measurement are not found in the simulation, however. A number of ripples appear on both the inside and the outside of the X-ray shell structure, and the authors attribute the ripples to nonuniformities, or hot spots, in the laser focus. In addition, one larger protuberance, perhaps $100\text{ }\mu\text{m}$ in size, was seen on the inner side of the shell. It seems likely that this protrusion is an indication of breakup of the pusher shell occurring during the laser pulse.

In a previous experiment, Ceglio and Coleman (*Phys. Rev. Lett.* **39**, 20; 1977) used a similar zone-plate, coded-aperture

technique to image alphaparticles from the thermonuclear burn region in an exploding pusher implosion at the Lawrence Livermore Laboratory Argus laser. They were able to detect burn regions smaller than $30\text{ }\mu\text{m}$ in targets which were initially about $90\text{ }\mu\text{m}$ in diameter, proving that the burn occurred near the target center, as was hoped, and not in the outer regions of the target.

A somewhat different approach to X-ray imaging of a laser fusion target was reported recently by Fenimore *et al.* (*Appl. Optics* **18**, 945; 1979). Their coded aperture was a plate with a large number of tiny pinholes. Since the signal-to-noise ratio is proportional to the square root of the number of holes, sensitivities comparable to that of the zone plate aperture were produced by choosing the appropriate aperture pattern. Fenimore was able to virtually eliminate undesired artifacts, such as "ghosts," in the reconstructed image. Furthermore, since digital (computer) reconstruction of the image was used, it was possible to avoid distortion of the image intensities relative to the source strength. Optical reconstruction tends to give an image for which isodensity contours correspond to isoemission contours of the source, but the image

density is not proportional to the source emission intensity. One drawback to using an aperture made with a large number of small holes is the difficulty of making each hole with an exactly known size, shape, and position to achieve the artifact-free images this method is capable of producing. Even though the 4×10^4 pinholes in the Fenimore aperture were $25\text{-}\mu\text{m}$ square, not as small as state-of-the-art fabrication techniques would permit, a very high-quality, linear image was obtained because of the care taken in the aperture construction. A camera aperture with $6\text{-}\mu\text{m}$ holes was recently built and should give increased resolution.

The technology for doing high-resolution imaging of particles and X-rays is advancing rapidly at present because of the need for improved diagnostics for laser-driven fusion implosions. The recently reported X-ray and alphaparticle imaging experiments have given valuable insight into the exploding-pusher implosion process, and we can expect further progress as high-resolution imaging becomes a routine diagnostic tool. It is quite likely that the imaging advances made because of laser-fusion needs will be of considerable use in other fields, particularly nuclear medicine. □

The present state of tranquility

from Leslie L. Iverson

MORE than 8000 tons of benzodiazepines were prescribed in 1977 in the United States. Some drugs in this group, such as diazepam (Valium) are used for their anti-anxiety properties, others such as flurazepam (Dalmane) and nitrazepam (Mogodon) as sleeping pills, and they have become the most widely used of all psychotropic drugs. It is only very recently, however, that progress has begun to be made in understanding how they work. This has been due to two important developments: the discovery of specific high affinity binding sites for radio-labelled benzodiazepines in brain (Squires & Braestrup *Nature* **266**, 732; 1977; Möhler & Okada *Science* **198**, 849; 1977), and the emergence of a plausible hypothesis to explain how benzodiazepines may interact with normal brain function. This hypothesis suggests that benzodiazepines act by enhancing the effectiveness of the important inhibitory chemical transmitter gamma-aminobutyric acid (GABA) in brain (for reviews see Iversen *Nature* **275**, 477; 1978; and Tallman *et al. Science* **207**, 276; 1980). With these two leads, biochemical research on benzodiazepine receptors in brain and their relation with GABA mechanisms has developed rapidly to become one of the most active areas of psychopharmacology.

A key question concerns the nature of the interaction between benzodiazepine and GABA receptors. The interaction is indirect, since the benzodiazepines do not simply mimic the actions of GABA, and

indeed do not compete directly with radioactive GABA for binding to GABA receptors in brain membrane preparations. Similarly GABA and related inhibitory compounds do not compete for the binding of radiolabelled benzodiazepine to their recognition sites. One possible clue about the nature of this interaction comes from a consideration of the mechanism by which GABA receptors control the excitability of brain cells. The effect of GABA when bound to its receptors is to trigger a specific change in the permeability of the cell membrane to chloride ions, and it is the movement of chloride which stabilises the electrical potential of the cell and reduces its excitability. The GABA receptor may thus be viewed as part of a macromolecular complex in which the GABA recognition site is associated with a chloride ionophore. Benzodiazepines might interact not with the GABA recognition sites but with the chloride channels to enhance the normal effects of GABA, as suggested by Guidotti *et al.* (*Nature* **275**, 553; 1978) and others. This view is also proposed by Simmonds (*Nature* **284**, 558; 1980) and supported by his new experimental findings. Simmonds studied the inhibitory effects of GABA in an *in vitro* preparation of small slices of rat cuneate nucleus using electrophysiological recording techniques. Addition of flurazepam to this preparation caused a small but significant potentiation of the effects of applied GABA. More importantly, flurazepam also significantly

Lynn R. Veaser is in the Physics Division, University of California, Los Alamos.

weakened the ability of the GABA blocking drug picrotoxin to antagonise the effects of GABA, while it had no effect on the GABA-blocking actions of another antagonist drug, bicuculline. The significance of this finding is that whereas bicuculline is thought to act by competing directly at GABA receptors, picrotoxin acts at a different site, perhaps related to the chloride ionophore. Although flurazepam and picrotoxin do not appear to compete for the same site, the result points again to an interaction of the benzodiazepine receptor with the chloride channels associated with GABA receptors. Simmonds suggests that the benzodiazepines might increase the duration of channel opening for each GABA receptor activation, or improve the efficiency with which GABA receptor occupation is associated with channel opening.

A similar conclusion has been drawn from other lines of research on GABA/benzodiazepine mechanisms. A number of laboratories have observed that GABA increases the affinity of binding sites for radiolabelled benzodiazepines in brain membrane preparations *in vitro* (Tallman *et al.* *Nature* **274**, 383; 1978; Briley & Langer *Eur. J. Pharmac.* **52**, 129; 1978; Martin & Candy *Neuropharmac* **17**, 993; 1978), and the latter authors also observed that this effect of GABA was considerably enhanced by chloride ions. Costa *et al.* (*Nature* **277**, 315; 1978) tested a number of other anions and claimed that their effectiveness in causing increases in the affinity of benzodiazepine binding correlated with their ability to replace chloride in GABA-mediated inhibitory events, although this conclusion has been challenged by Martin and Candy (*Neuropharmac.* **19**, 175; 1980).

An unexplained feature of the GABA effect on benzodiazepine binding is that a number of compounds which mimic the inhibitory actions of GABA on neuronal firing and which compete for ³H-GABA binding sites fail to enhance benzodiazepine binding (Karobath *et al.* *Nature* **278**, 748; 1979 and some of these compounds may even antagonise the effects of GABA on benzodiazepine binding (Braestrup *et al.* *Nature*, **280**, 331; 1979). This suggests that the GABA recognition site which mediates changes in benzodiazepine receptor affinity may differ in specificity from that of the normal GABA inhibitory receptor.

Regardless of its pharmacological characteristics, however, the existence of a GABA recognition site which can influence the binding properties of the benzodiazepine receptor adds further support to the GABA/benzodiazepine interaction hypothesis, and an interesting recent observation is that GABA continues to exert a modifying action on the affinity of benzodiazepine binding even in solubilized

benzodiazepine receptor preparations (Gavish & Snyder *Life Sci.* **26**, 579; 1979).

Our understanding of the complexity of GABA mechanisms in brain has been increased by the finding of an apparently novel type of GABA receptor which has a different drug specificity from the conventional post-synaptic inhibitory receptors for GABA. Bowery *et al.* (*Nature* **283**, 92; 1980) have summarised the evidence for the existence of such a novel GABA receptor. GABA at micromolar concentrations inhibits the release of monoamine transmitters from nerve terminals in the brain and in the peripheral autonomic nervous system, but unlike conventional GABA responses these effects are not antagonised by bicuculline. The anti-spastic drug baclofen (8-parachlorophenyl GABA) potently mimics these presynaptic effects of GABA, although this compound is inactive on post-synaptic GABA receptors.

Furthermore, the presynaptic GABA sites fail to recognise compounds such as 3-aminopropanesulphonic acid and muscimol, which are potent GABA-mimetics on post-synaptic GABA receptors. These observations suggest that a different type of GABA receptor may be involved, bringing GABA into line with other CNS neurotransmitters for which multiple receptor types are now thought to exist.

Research on the minor tranquilisers and brain GABA mechanisms is far from tranquil, although from this turmoil a better understanding is emerging of the cellular and molecular pharmacology of this major group of psychotropic drugs. At the same time rapid developments are now taking place in isolating and purifying benzodiazepine receptors and identifying possible naturally occurring ligands for them. □

Cataclysmic conferences show convergence

from Robert P. Harkness and Brian McMahon.

SINCE the discovery some sixteen years ago that novae and related variable stars were binary systems, theoreticians and observers have devoted a great deal of effort to understanding how the interactions between the two stellar components give rise to energetic outbursts on timescales ranging from weeks to thousands of years. Two recent informal conferences* on the nature and properties of cataclysmic variables have given a timely glimpse of the state of the art.

Our current understanding of dwarf novae was emphasised at Cambridge. These systems typically undergo irregularly spaced outbursts of some 2-4 magnitudes every few weeks, lasting for a few days at a time. It is believed that the energy required to power the outbursts comes from gravitational energy of matter which is transferred from the binary companion star to a white dwarf primary, and that this transferred matter forms a disc around the primary which is stable between outbursts. J. Pringle (Cambridge) presented a series of models due to him and G. T. Bath (Oxford) which follow the evolution of accretion discs following the injection of material from the companion star, demonstrating how the density and temperature distribution and the timescale of the resulting structural changes depend on the as yet poorly understood "viscosity" of the disc.

The existence of such accretion discs is rapidly being confirmed, particularly from combined ultraviolet and optical studies with the International Ultraviolet Explorer satellite (IUE) and ground-based optical telescopes. J. Whelan (Cambridge) showed the continuum spectra covering the wavelength region 1000 — 7500 Å of three dwarf novae taken simultaneously in the

optical and ultraviolet, and compared them with the spectral energy distribution predicted for a steady-state, optically thick accretion disc. For the two systems VW Hyi and EX Hya, no single stellar-type continuum could fit the data, but the fit was very reasonable for simple discs, and allowed disc effective temperatures to be deduced. The third system, BV Cen, showed a different flux distribution which was thought to yield information about the secondary star.

J. Papaloizou (Cambridge) discussed systems of the SU UMa type, which occasionally show more luminous "superoutbursts" lasting about a week. During these eruptions, the light curve contains periodic humps with a period differing slightly from the binary orbital period of the system. It is thought that these so-called "superhumps" may be due to a tidal process which excites toroidal modes of oscillation in the secondary and thus modulates the mass transfer flow with waves of the same period.

Such "superhumps" were observed by M. Friedjung (Paris) in the 1978 outburst of the recurrent nova WZ Sge. Spectroscopy with IUE indicates a continuum consistent with an accretion disc spectrum rather than the spectrum typical of a classical nova, but double-peaked emission lines which become narrower at outburst are interpreted as features representing an outer, slowly-rotating disc. Such a structure was suggested as an alternative cause of "superhumps".

At Austin, much new observational work on dwarf novae was presented, all of which was interpreted in terms of the now standard

Robert P. Harkness and Brian McMahon are in the Department of Astrophysics, University of Oxford.

Leslie L. Iversen is Director of the MRC Neurochemical Pharmacology Unit, Department of Pharmacology, Medical School, Cambridge.

accretion disc model. High-quality optical spectroscopy of the system SS Cyg was reported by E. Nather (Austin) and also by A. Cowley (Dominion Astrophysical Observatory), leading to a revised period of 0.275 days, but yielding no definite results for the mass ratio of the components. Abundances in the accretion disc were investigated by S. Wyckoff (Phoenix), who reported an excess of helium over hydrogen in the disc of EX Hya.

Satellites have produced much new data, and ultraviolet observations from IUE were reported from P. Szkody (Washington) and also by G. Fabbiano (Harvard-Smithsonian Centre for Astrophysics). F. Cordova (Los Alamos) presented recent X-ray results from the Einstein satellite for a number of systems, all of which had significant hard X-ray fluxes, while no ultrasoft emission at energies around 50 eV (as previously seen at outbursts of SS Cyg and U Gem) was detected. The failure of previous satellites to observe X-ray emission from most cataclysmic variables allows an upper limit to be put on the mean X-ray luminosity of 10^{31} ergs s^{-1} . Most of the ultraviolet and X-ray spectra presented were in reasonable accord with the general predictions of the accretion disc model.

Einstein and IUE observations of nova-like variables were presented by E. Sion (Villanova). These objects show a wider range of properties than ordinary dwarf novae. V471 Tau, in the Hyades cluster, has a primary which may be the hottest white dwarf known, with a temperature of over 60 000 K. High-excitation spectral lines of silicon and carbon seen in absorption are thought to arise in the photosphere of the white dwarf, and there is no evidence for an appreciable accretion disc. TT Ari and CD-42°4462 display similar features, but the latter in outburst has a very similar spectrum to that of the dwarf novae.

At Cambridge, an attempt to model the colours and eclipses of the nova-like variable RW Tri was made by J. Frank (Leicester) and co-workers. They assumed the presence of an accretion disc and a secondary on or near the main sequence. If their model is correct, it implies a mass for the white dwarf component of 1.4 times the mass of the sun, which is uncomfortably close to the Chandrasekhar maximum. This result was challenged in discussion by J. Smak (Warsaw), who claimed that radial velocity measurements on the system give a figure much closer to 0.4 solar masses.

The properties of systems related to the short-period binary AM Her were reviewed at Austin by J. Liebert (Tucson). These are characterised by strong linear and circular polarisation, and are usually X-ray sources. They exhibit no nova-like eruptions, but do have distinct "low" and "high" states. They are now usually interpreted as systems containing an accreting highly-magnetised white dwarf.

It is conjectured that the high magnetic fields disrupt or prevent a disc from forming, so that accretion columns are produced at the magnetic poles of the primary. Four of these systems are now known, the most recent discovery being the X-ray source 2A0311-227. Optical and X-ray light curves for this 81-minute eclipsing binary were presented by J. Patterson (Minnesota). The X-ray flux has a soft black body component, and a hard bremsstrahlung component with energies in excess of 11 keV. This latter component is not eclipsed at all, while the soft X-ray flux is halved at eclipse. This was interpreted as the eclipse of a hot spot in the accretion stream by the accretion column itself. This same object was discussed at Cambridge, where M. Watson (Leicester) explained the narrow spike appearing in one of the eclipse components of the optical light curve by the passage of the secondary star across the line of sight. The other eclipse features are then due to the eclipse of the magnetic poles by the body of the white dwarf itself.

While our knowledge of dwarf novae and related systems appears to be increasing steadily, less progress is being made on the problem of the classical novae, which undergo outbursts perhaps only every 30 000 years, but with ranges of 8-17 magnitudes. It is thought that these, too, are binary systems, but it is argued that the violence of their outbursts must be explained by thermonuclear ignition of hydrogen-rich material accreted by the white dwarf. It is generally agreed that the nova outburst produces an optically thick envelope around the whole binary system; as this envelope thins, one sees deeper into regions of higher temperature, so that the flux maximum moves into regions of shorter wavelength. Thus the bolometric (total) luminosity remains essentially

constant for some time, while the visual luminosity drops off rapidly. If this picture is right, one would expect to see enhanced abundances of such elements as carbon and nitrogen which are necessary to fuel the thermonuclear runaway. At Cambridge, C. Penn (London) discussed ultraviolet observations of Nova Cyg 1978 following outburst, the first time a classical nova eruption has been seen by IUE. The development of several spectral lines was followed for nearly a year, and in the early stages of the outburst many showed the P Cygni profiles expected for outflowing material. Measurements were made of carbon and nitrogen which showed them to be highly overabundant in the nebular phase.

The Austin meeting heard from M. Shara (Montreal) a theoretical paper seeking to derive the empirical relation which has long been established between the maximum magnitude of a classical nova and the speed with which it subsequently drops in luminosity by three magnitudes. This relation is of wider importance in astronomy, since a good theoretical basis for the empirically observed result would allow classical novae to be used as "standard candles" and distance indicators with greater confidence.

The Austin meeting was the fifth in an irregular series of workshops on cataclysmic variables, the Cambridge meeting the first such to be held in Britain. The range of activity demonstrated the excitement and dynamism of the subject. Now that interpretations seem to be hardening on specific theoretical models, a flurry of more detailed calculations can be anticipated, including as a major contribution models of accretion disc structure that can be compared with a large array of observational tests. □

Structural aspects of recognition and assembly in biological macromolecules

from a Correspondent

THE last decade has witnessed a dramatic increase in our understanding of the structural, thermodynamic and kinetic factors which allow recognition and assembly in molecular biology. More complex supramolecular systems have been subjected to the techniques of amino acid and nucleic acid sequencing, to high resolution X-ray analysis and spectroscopic probes, and to mechanistic and functional investigations. We can now begin to assess the role of shape, hydrophobicity, flexibility and other factors in the processes of recognition between biological molecules. This was the theme of the VIIth Aharon Katzir-

Katchalsky Conference held at Kibbutz Nof Ginossar on the shore of the Sea of Galilee, Israel, earlier this year.

Of the many biomolecular systems, enzyme substrate interactions have been most thoroughly investigated; but only recently have very high resolution (~ 1.5 Å) X-ray studies allowed precise definition of the molecular geometry and thermal vibrations — a measure of at least one kind of flexibility. This is proving a critical testing ground for older hypotheses of catalytic mechanism involving steric strain, charge relay systems and covalent intermediates. For instance, M.N. James and his colleagues (University of Alberta) have

*An Informal Conference on Cataclysmic Variables, held at the Institute of Astronomy, Cambridge, England, March 20-21, 1980; and The Fifth Workshop on Cataclysmic Variables and Related Objects, Austin, Texas, March 24-26, 1980.

cast doubt on the accepted mechanism of serine proteinases by their detailed analyses of the serine proteinase A from *Streptomyces griseus* where they find bond distances which are considerably longer than expected in the charge relay system, tetrahedral intermediate model now described in all our biochemistry text books. They also demonstrate the existence of highly ordered water in the active site which must be displaced by the solvent; the release of this water would be entropically favourable to the assembly of the complex. Another striking feature of this work is the observation that although there is high thermal motion of the "jaws" of the substrate-binding site in the native enzyme, the polypeptide in this region becomes significantly more rigid in the presence of a ligand. The flexibility may play an important role in the docking of substrate with the enzyme. Similar ordering is found by R. Huber's group (Max-Planck-Institute, Munich) in a disordered "activation domain" in trypsinogen which includes part of the specificity site when the bovine pancreatic trypsin inhibitor or other inhibitors are bound in the presence of an activator dipeptide. K. Wüthrich (ETH, Zürich) reported similar order-disorder transitions in solution using nuclear magnetic resonance studies, and it is now becoming clear that such transitions may be of general functional importance in interactions between different biological systems.

Progress is also being made in our understanding of the assembly of fibrous proteins, especially in muscle components. For instance, the amino acid sequence of tropomyosin shows a number of periodic and non-periodic features which can be correlated with its two-stranded coiled-coil structure, with its head-to-tail aggregation into long filaments, and with its interactions with F-actin and troponin (L.B. Smillie, University of Alberta). A mechanism was described whereby tropomyosin regulates muscular contraction by responding to structural changes induced in troponin in Ca^{++} ions and consequently rolling across the surface of actin, thereby uncovering the site for actin binding to myosin (D.A.D. Parry, Massey University). New results concerning the binding of myosin cross-bridges to actin suggest that these may not change their slope during contraction (K. Holmes, Max-Planck-Institute, Heidelberg) and that the contractile force may result from a helix-coil transition in the S_2 fragment of myosin (W.F. Harrington, Johns Hopkins).

Although models for the recognition of DNA by proteins have until recently assumed a dimorphic DNA consisting of only the B and A forms, it is now clear that the DNA molecule is much less rigid than previously believed. S. Arnott (Purdue University) presented evidence from fibre X-ray diffraction studies for a great variety of possible DNA double helical structures.

These structures are mostly right-handed, but specific polymers of alternating purine and pyrimidine bases were found to form left-handed duplexes similar to the Z DNA structure (*Nature* 282, 680; 1979). Although it is generally believed that most DNA has a classical DNA double helix most of the time it is probable that such specific structures in parts of DNA may be involved in their recognition by binding proteins. H. Sobell (University of Rochester) suggested that thermally induced low-frequency-mode oscillations in a DNA duplex structure may play a central role in transitions between different structures and in the binding of other molecules to DNA.

The exact pitch of DNA in a natural environment may also differ from that in the fibres which formed the basis of the Watson-Crick double helix. A. Klug (MRC, Cambridge) described some recent experiments aimed at measuring the true pitch of DNA in solution by cleaving molecules on a very flat mica surface with DNase I. The pitch was found to be 10.6 bases; significantly different from the 10.0 bases measured in X-ray fibre studies, but quite similar to the average value of 10.4 suggested from DNase I digestion of chromatin DNA. E. Trifonov (Weizmann Institute) described a quite surprising result that the distributions of the dinucleotides CG, TG, TA and TT along DNA sequences found in chromatin are periodic. The period is about 10.5 bases which agrees within error limits with previous estimates of the pitch of DNA in chromatin.

Evidence for small variations in t-RNA structures which may be important in their recognition is also forthcoming from the structure analyses of yeast initiator t-RNA^{Met} by J. Sussman (Weizmann Institute; see *Nature* 278, 188; 1979) and yeast t-RNA^{Asp} by J.-C. Thierry (University Louis Pasteur, Strasbourg). In contrast to the class I, chain elongation t-RNA^{Phe}, the t-RNA^{Met} is a chain initiation t-RNA and the t-RNA^{Asp} is a class II t-RNA. Both show the canonical "L" shaped conformation of t-RNA^{Phe} but there are differences in the closing of the angle between the two long helical stems of the "L" by about 8° in t-RNA^{Met} and the opening of this angle by about 10° in t-RNA^{Asp}. P. Sigler (Chicago University) proposed an intriguing hypothesis for the molecular basis of the unique role played by eukaryotic initiator t-RNAs in protein synthesis, suggested to him by the t-RNA^{Met} crystal structure. The characteristic sequence features which are common and unique to eukaryotic initiator t-RNAs are clustered in a 10Å domain at the union of the D and T loops and fold in a way that preserved the backbone trace seen in other t-RNAs. However, the potential for unique interactions with cognate proteins and nucleic acids makes this region a likely focus for the specific interactions with the ribosome, that underlie the special role of t-RNA^{Met}. These

interactions stabilise a factor-mediated 80s ribosomal initiation complex which recognises the first AUG encountered from the 5'-terminus of the m-RNA as the "Start" signal.

The recently determined structure of the Southern Bean Mosaic Virus, SBMV (M. Rossmann, Purdue University) shows that there may be common features in the assembly of spherical viruses. A comparison of its structure with that of Tomato Bushy Stunt Virus, TBSV (S. Harrison, Harvard University) indicates a conformation similar to that of the shell domain of the latter. The SBMV protein also uses a similar pattern of contacts with neighbouring molecules within the viral capsid. Studying the inter-subunit contacts in the shell domain of TBSV, Harrison has found an alternating pattern of polar and hydrophobic patches, as previously found for the lateral interactions in Tobacco Mosaic Virus (A. Bloomer *et al.*, MRC, Cambridge), and he suggested that this may be a general feature of interfaces which can exist in more than one quasi-equivalent state. The three types of interaction between hexons in the capsid of adenovirus (R.M. Burnett, Biocentrum, Basel) can now be understood in terms of the pseudo-hexagonal symmetry of the trimeric protein. The symmetry facilitates packing, whereas the departures from it allow the different interactions.

It is hoped that the proper understanding of recognition in self-assembling systems may provide the basis for a contribution of molecular biology to medicine. X-ray studies of fibres and crystals of deoxygenated sickle-cell haemoglobin have led to a model for the interactions specific to the mutant haemoglobin. M. Gorecki (Weizmann Institute) described the design of small peptides and esters of hydrophobic amino acids based on this model. These compounds appear to inhibit sickle cell haemoglobin aggregation, they permeate erythrocytes rapidly, and thus appear to have considerable potential for the therapeutic treatment of sickle cell anaemia. In a similar way analyses of three-dimensional models for insulin and glucagon receptor interactions (T.L. Blundell, Birkbeck College, London) have led to pharmaceutical activity in the design of insulin agonists and glucagon antagonists which may be of value in the treatment of diabetes. The recent determination of the first three-dimensional structure of a long neurotoxin by M. Walkinshaw and W. Saenger (Max-Planck-Institute, Göttingen) has suggested a refined model for the multipoint attachment of toxins to the nicotinic acetylcholine receptors which must stimulate the design of new neuromuscular blocking agents. Even greater rewards may come from studying some of the more complex structures. For instance, it will be surprising if our new knowledge of virus coat structures does not lead to the design of some useful antiviral agents. □

ARTICLES

Northward displacement of north-central British Columbia

J. W. H. Monger

Geological Survey of Canada, 100 West Pender Street, Vancouver, Canada V6B 1R8

E. Irving

Earth Physics Branch, Energy, Mines and Resources, Ottawa, Canada K1A 0E4

An extensive terrain in north-central British Columbia, apparently moved ~1,300 km northwards along the western fringe of North America during the late Cretaceous or early Tertiary, perhaps in much the same fashion as Baja California is doing today. Large anticlockwise rotations occurred within this terrain.

THE western, volcanosedimentary 'eugeosyncline' of the North American Cordillera is a collage of terrains of varied origins that apparently became attached to the western continental margin of the continent at different times¹⁻³. These terrains now lie to the west of the line WW in Fig. 1, which is the probable western limit of the ancient North American craton. It is delineated by the westernmost limit of Palaeozoic and Precambrian strata whose continuity with the craton can be reasonably inferred, and by the transition of the ⁸⁷Sr/⁸⁶Sr initial ratios from <0.704 to the west (indicating that the underlying crust is of oceanic or island arc origin) to >0.706 to the east (indicative of continental crust)⁴. EE is the eastern limit of Cordilleran deformation.

The concept of a tectonic collage derives from several lines of evidence. Each terrain has an early Mesozoic and late Palaeozoic stratigraphy which differs from that of contiguous terrains^{5,6}, each has early Mesozoic and late Palaeozoic faunas that differ from coeval faunas along the ancient cratonic margin^{7,8}, and several have yielded palaeomagnetic results indicating that they have been displaced latitudinally or rotated by large angles about local vertical axes⁹⁻¹². The Alaskan peninsula and the Chugach Mountains (SAJ of Fig. 1), composed of Jurassic and younger rocks, have been displaced 1,500 km from the south since the Jurassic¹⁰. A western terrain¹³, sometimes called Wrangellia⁵, with distinctive Triassic stratigraphy, comprises parts of southeastern Alaska, Queen Charlotte Islands, Vancouver Island and the Hell's Canyon area of eastern Oregon the latter denoted by '?' in Fig. 1b. Wrangellia has moved at least 1,500 km, and possibly as much as 5,000 km, from the south since the late Triassic^{9,11}. Much of southeastern Alaska is underlain by the Alexander terrain, with characteristic Palaeozoic and Triassic strata¹⁴. Recent palaeomagnetic results¹⁵ from the Alexander terrain indicate large anticlockwise rotations and yield Triassic palaeolatitudes consistent with those for the craton. Instances of northward displacements and clockwise rotations have been reported from batholithic rocks^{16,17}. The Coast Range of Oregon^{18,19}, and the Transverse Ranges of southern California²⁰ have rotated clockwise by ~60° during the Tertiary. We now present palaeomagnetic evidence from a reconnaissance survey for the northwards displacement and differential rotation of an extensive tract of country in northern British Columbia known as the Stikine terrain, together with rocks of the Cache Creek Group that lie along its eastern border.

The integrity of the Stikine terrain during the early Mesozoic is shown by the distinctive late Triassic and early to

middle Jurassic rocks that extend across it. It is bordered to the east by the Cache Creek Group, which lies just west of WW. The Cache Creek Group is interpreted as late Palaeozoic-early Mesozoic ancestral Pacific Ocean crust with its overlying deposits that were thrust westwards on to the Stikine terrain in post-Oxfordian time^{2,6,21}. To the east of WW there are volcanogenic terrains that may have formed along or immediately offshore from the Pacific margin of North America during the Mesozoic². The displacement of terrains in the western Cordillera from the south has been suggested first to account for the juxtaposition of warmer-water western Triassic faunas and cooler-water faunas in the Rocky Mountains⁷, and second to explain the peculiar distribution of Permian fusulinids whereby the Cache Creek Group, with its distinctive verbeekiniid or Tethyan faunas, was inserted between schwageriniid or American faunas that occur to the west in the Stikine terrain and to the east and south in south-central British Columbia, Oregon and California⁸. Our palaeomagnetic studies were undertaken to make quantitative estimates of any displacements that might have occurred.

Palaeomagnetic results

The rocks sampled are Upper Triassic and Lower Jurassic volcanogenic strata, typical of the Stikine terrain, and a Lower Cretaceous layered intrusion cutting deformed Cache Creek rocks (Fig. 2). These three rock units were sampled at 42 sites distributed among six localities. Within each locality the sequences are continuous, with no field evidence of relative motion. Attitudes of bedding, flow layering or compositional layering were measured at each site, with the exception of some massive augite porphyries at the base of the Triassic section (Table 1, section (7)), but their continuity with overlying, compositionally similar, uniformly dipping flows suggests they have the same attitudes. Three to six oriented samples (cores or hand samples) were obtained from each site. Two cores were cut from each hand sample, making a total of 236 cores. Two specimens were cut and measured from each core.

The oldest rocks studied are from the Savage Mountain Formation of the Takla Group of late Carnian to earliest Norian age²². The Savage Mountain Formation consists of up to 5,000 m of dark grey-green, locally red, pillow basalt and massive subaerial flows, with associated volcanoclastics. Samples were obtained from two localities, mainly from subaerial flows. At all collecting sites, with the exception noted above, attitudes could be accurately measured. When sampling

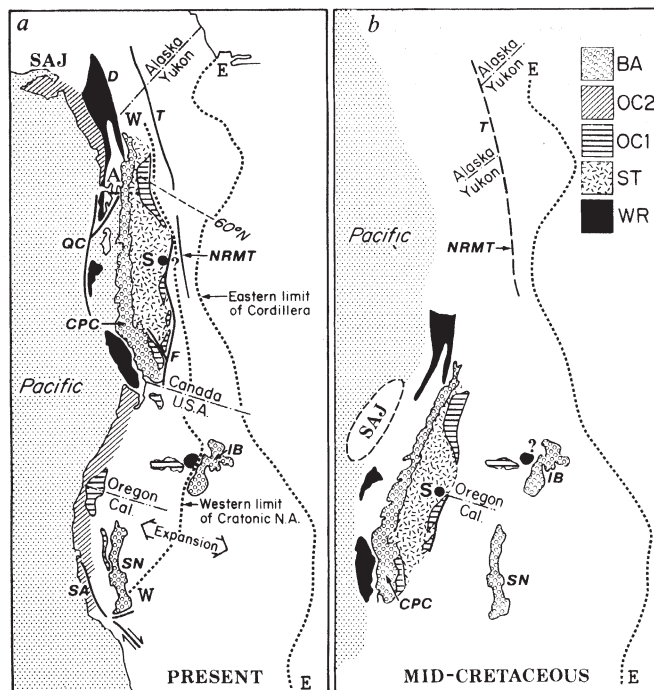


Fig. 1 *a*, A sketch-map showing some of the major tectonic elements of the Cordillera. EE is the easterly limit of deformation; WW, the westerly limit of the North American craton. BA, major plutons (CPC, Coast Plutonic Complex; IB, Idaho batholith; SN, Sierra Nevada batholith); WR, Wrangellia; ST, Stikine terrain; OC1, mid-Jurassic and older oceanic crust and rocks deposited on it; OC2, late Mesozoic and Tertiary oceanic crust and rocks deposited on it; A, the Alexander terrain is the blank area. *b*, The proposed position of the Stikine and Wrangellian terrains during the late Triassic to mid-Cretaceous. The Tertiary expansion of the Basin and Range Province is restored, and Wrangellia and the Coast Plutonic Complex are attached to the Stikine terrain in their present relative positions. These terrains are moved undistorted. We consider only gross latitudinal displacement, and no attempt is made to restore the original geometries of individual elements of the Stikine and Wrangellian terrains. Certainly large internal rotations are involved (see text). Major right-lateral strike-slip faults are marked: D, Denali; F, Frazer; NRMT, Northern Rocky Mountain Trench; QC, Queen Charlotte Island-Chatham Strait system; SA, San Andreas; T, Tintina. S is the sampling area, and SAJ is the late Jurassic displaced terrain of southern Alaska¹⁰. In *b* the Alaska-Yukon border is displaced on the Tintina-Northern Rocky Mountain Trench lineament to show the amount of motion that can be demonstrated geologically on one of the fault systems that might record some of the northward displacement of the Stikine terrain. The total displacement occurred by motions of comparable magnitude distributed among several such faults (see text).

Takla and Hazelton rocks, preference was given to deuterically oxidized basalts, because such rock types have been shown to be very stable magnetically²³, and our preliminary studies confirmed that this was also true for the sequences we have sampled (Fig. 3). Such oxidized rocks can usually be recognized by diagnostic minerals such as iddingsite in olivine basalts, because they are often reddish, and because they yield grey-red mud when drilled. The Takla basalts have alkaline affinities, and are generally porphyritic, with abundant clinopyroxene, feldspar, and altered olivine phenocrysts. The directions of natural remanent magnetizations (NRM) at each site are sometimes scattered, but are more commonly well grouped. During cleaning by alternating field (30–80 mT) or thermally (550–600 °C), they coalesce into good groups. The same directions persist after treatment in fields of 150 mT, and in some specimens, up to 200 mT. Blocking temperatures of this high coercive force magnetization are in the range 500–650 °C, and haematite and magnetite blocking temperatures often occur in the same specimen (Fig. 3). The directions remain unchanged between 500 and 650 °C, the haematite and magnetite magnetizations having identical directions. After cleaning, and after correction is made for the dip of the beds, the directions are westerly. This direction characterizes the

Savage Mountain Formation in the places sampled. The two localities have statistically identical inclinations, but their mean declinations differ by 19° as if they have been rotated relative to one another about vertical axes (Fig. 4, Table 1). The mean inclination of 41° yields a palaeolatitude of 23°, which is significantly lower than the late Triassic palaeolatitudes calculated from cratonic North America (Fig. 5).

Lower Jurassic rocks of the Hazelton Group were sampled at three localities spread over 100 km (Fig. 2). The Hazelton Group consist of calc-alkaline volcanics and intercalated sediments. It ranges in age from late Sinemurian to early Callovian, and extends across the eastern and southern parts of the Stikine terrain²⁴. The group contains three formations: Telkwa (bottom), Nilkitkwa and Smithers (top). Late Sinemurian subaerial red basalt, and red fine-grained tuff of the Telkwa Formation were sampled from the Red Creek–Two Lake Creek area, and the Driftwood Range. Toarcian basalt and red tuff of the Nilkitkwa Formation were obtained from the Bait Range. The directions of NRM at each collecting site are usually grouped together but are occasionally scattered. Cleaning by alternating fields and heating yields well-grouped directions. The coercivity commonly exceeds 100 mT, and blocking temperatures are in the range of 500 to 700 °C, except in some tuffs in which the range is 200–600 °C (Fig. 3). The directions at the sites within a locality agree reasonably well. Moreover, the mean inclinations are statistically identical from locality to locality, but the declinations range from south-west to north-northeast (Fig. 4, Table 1), as if the localities have been differentially rotated by large angles about vertical axes. Alternatively, one may argue that these directions (unrotated) could be the resultants of unresolved antiparallel magnetization acquired over a substantial time interval during lithification when the geomagnetic field reversed in direction. This is improbable

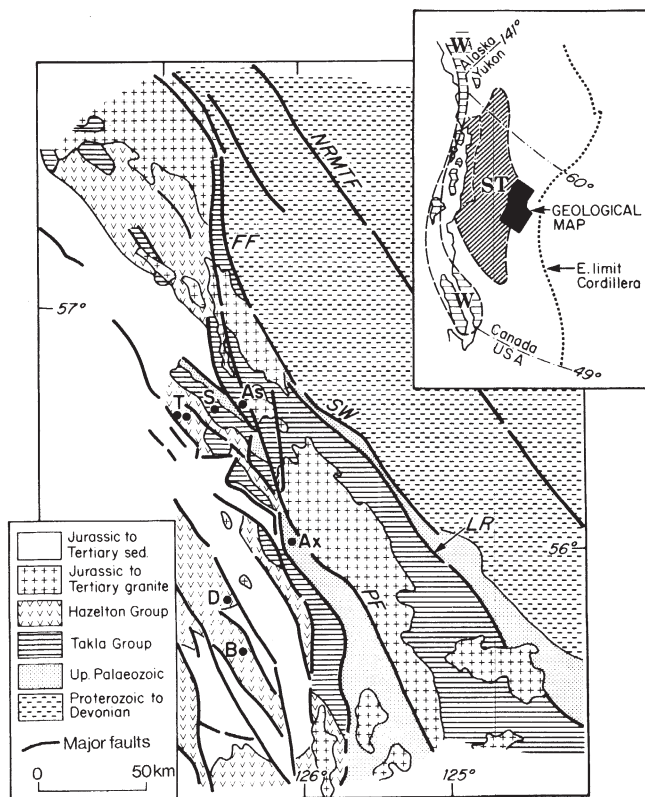
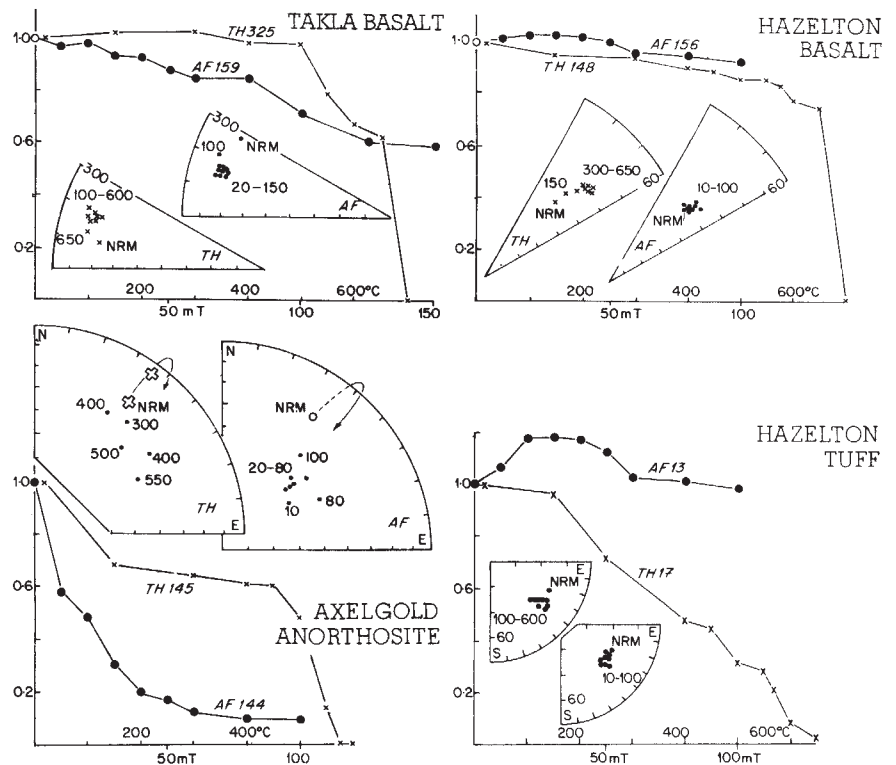


Fig. 2 Geology of the sampling region. The sampling localities are: As, Asitka Peak; Ax, Axelgold intrusion; B, Bait Range; S, Sustut Peak; and T, Two Lake Creek and Red Creek localities. The inset shows the general location and the two major displaced terrains, Wrangellia, W, and the Stikine terrain, ST. The faults marked are: FF, Findlay; LR, Lay Range; NRMT, Northern Rocky Mountain Trench; PF, Pinchi; and SW, Swannell.

Fig. 3 Demagnetization studies. Examples of thermal (TH) and alternating field (AF) demagnetization of pairs of specimens from the same oriented core. Directions are plotted on equal-area projections, the plane is the present horizontal. The numbers on the demagnetization graphs are the intensities of natural remanent magnetization (NRM) in units of 10^{-3} A m^{-1} . The Takla and Hazelton basalts are deuterically oxidized; the directions are little changed throughout the demagnetization range; remanent coercive forces are high; blocking temperatures appropriate to haematite and fine-grained magnetite occur; both minerals can be identified in polished section and both have the same magnetization direction; inclinations are positive (downward). In the Axelgold anorthosite the directions initially have negative inclination and they migrate rapidly to the characteristic direction (inclination positive) for the intrusion during removal of a soft magnetization which is randomly directed from core to core and which we attribute to lightning; there is a well-marked 'magnetite' blocking temperature; in polished section magnetite occurs as very fine rods in felspar crystals; remanent coercive forces are typical of those of fine-grained magnetite; haematite magnetizations do not occur. The red tuff from the Hazelton Group has negative inclination and the direction is little changed through the demagnetization range; the magnetization is due to very fine-grained hematite readily observed in polished section.



because the haematite and magnetite magnetizations that probably formed at somewhat different times always have identical directions (Fig. 3). Moreover, it is highly improbable that such a complex magnetic structure would have produced inclinations of both normal and reversed magnetizations that are consistent from locality to locality. A further possibility is that the directions (unrotated) are the resultants of primary magnetizations and stable secondary magnetizations, the latter acquired long after formation. Again, such an interpretation is highly improbable for the two reasons just given, and because we have been careful never to sample close to later igneous bodies or zones of secondary mineralization. Therefore we conclude that large relative rotations have occurred. A mean inclination (54°) is obtained by rotating all site directions so that the mean declinations at each locality agree (Hazelton rotated, Fig. 4). The mean palaeolatitude of 35° is 14° farther south than expected, but the difference is marginally significant. The large error of Fig. 5 arises from the considerable uncertainty in the early Jurassic field direction for cratonic North America as it is presently known (see below). These results from the Hazelton Group provide one of the few well-documented instances of reversals of magnetization in early Jurassic rocks.

The third unit is the Axelgold Intrusion, a layered body, 12 by 4.5 km in extent²⁵. Its age is less well defined than in the previous two units. The body intrudes highly deformed rocks of the Cache Creek Group, and cuts structures that, 25 km to the north, are congruent with structures of post-Oxfordian age. The intrusion is therefore post-Oxfordian. Two biotites and one hornblende from the reaction rim of a granite xenolith in the interior of the intrusion, and from a micaceous gabbro close to the contact, yield K-Ar ages in the range 98–108 Myr (ref. 25). A second hornblende gives a K-Ar age of 155 Myr, but is considered²⁵ to "reflect the presence of excess argon". A mid-Cretaceous age is the most probable²⁵, but an early Cretaceous age is also consistent with the fact that we have found only normal polarities in the body, and normal polarity typifies the geomagnetic field in the interval 82–110 Myr (ref. 26). If it were late Jurassic or early Cretaceous, both normal and reversed polarity would have been expected in such a

large slowly-cooled body, because field reversals were then frequent²⁶.

The intrusion comprises alternating layers of anorthosite and anorthositic gabbro, ranging from ~10 to 50 cm in thickness, and dipping 10 – 15° . Anorthositic samples were chosen preferentially because preliminary studies showed them to be more stable magnetically. The layering was presumably produced by gravity-controlled magnetic currents and is therefore considered to be a good indicator of the palaeohorizontal. The NRM directions are usually scattered, but after cleaning in alternating magnetic fields they become well grouped, the characteristic magnetization having north-northeast declinations and steep downward inclinations (Fig. 4). The coercivity commonly exceeds 50 mT and the blocking temperatures are in the range of 500 to 600 °C. The mean inclination relative to layering is 69° , yielding a palaeolatitude of 52° , about 13° lower than expected for the mid-Cretaceous. It is possible that the layering was not formed horizontally, or that tilting occurred before the acquisition of remanence whilst the body was hot, and this estimate of palaeolatitude would then be in error. Note that the directions of magnetization relative to present horizontal (Table 1) yield a palaeolatitude of 40° which is even more aberrant. A further source of error is in the age of the intrusion. If it were late Jurassic, then the estimate of palaeolatitude would be in good agreement with that expected for North America at that time (see ref. 12, Fig. 15), implying that both the Cache Creek terrain, into which the body is intruded and the Stikine terrain were then in place relative to cratonic North America. Thus it is important to note that our preferred interpretation of Fig. 5 is uncertain because of the difficulty in determining accurately the age and the palaeohorizontal.

Discussion

The results for these three time intervals all yield mean palaeolatitudes that are close to the latitude of the California–Oregon boundary. The Stikine terrain is geologically a coherent entity so that although the observations are from a geographically restricted area they, nevertheless, indicate that from the late Triassic until sometime in the Jurassic or

Table 1 Summary of results

	<i>Th</i>	<i>B</i>	<i>n</i>	<i>D, I</i>	<i>k</i>	α_{95}°	Pole λ, ϕ	(dm°, dp°)
(1) Axelgold, 98–108 Myr	270	13	67	025°, +69°	70	05	76°N, 033°W	(08, 07)
(2) Axelgold, uncorrected	270	13	67	032°, +60°	92	04	64°N, 012°W	(07, 05)
(3) Hazelton, mid-Toarcian	170	07	37	359°, +55°	15	16	70°N, 057°E	(22, 16)
(4) Hazelton, Sinemurian	070	04	28	242°, +56°	29	18	17°N, 175°W	(25, 18)
(5) Hazelton, Sinemurian	125	04	20	114°, -52°	14	25	40°N, 144°E	(35, 24)
(6) Takla, late Triassic	150	06	42	300°, +44°	31	06	38°N, 133°E	(15, 10)
(7) Takla, late Triassic	200	08	42	281°, +38°	18	07	24°N, 146°E	(16, 09)

(1) Axelgold layered intrusion, Axelgold Range 56.2°N, 126.1°W, cleaning 30–60 mT, normal polarity, direction relative to layering. Axelgold, directions relative to present horizontal. The precisions of (1) and (2) do not differ significantly. (3) Nilkitwa Formation, Hazelton Group, Bait Range 55.6°N, 126.4°W, cleaning 550–600 °C, two sites reversed five sites normal polarity. (4) Telkwa Formation, Hazelton Group, Driftwood Range 55.8°N, 126.6°W, cleaning 80 mT and 550 °C, all normal polarity. (5) Telkwa Formation, Hazelton Group, Red Creek–Two Lake Creek area 56.5°N, 126.8°W, cleaning 550–600 °C and 60 mT, all reversed. (6) Savage Mountain Formation, Asitka Peak 56.7°N, 126.4°W, cleaning 60 mT and 500–550 °C, all normal polarity. (7) Savage Mountain Formation, Sustut Peak 56.6°N, 126.5°W, cleaning 60 mT and 500–550 °C, all normal. Directions in (3) to (7) are relative to bedding. *Th*, thickness of the section sampled in metres; *B*, number of sites each given unit weight in the analysis; *n*, number of cores (two specimens per core studied); *D, I* the mean direction; *k*, Fisher's estimate of precision; α_{95} , the 95% circle of confidence³⁵. The position of the north pole is given; λ latitude, ϕ longitude. Unit weight given to each sampling site in the analysis.

Cretaceous the Stikine terrain was located ~1,300 km south of its present position relative to North America (Fig. 1).

The agreement between the mean values is excellent, better than would be expected from the associated errors (Fig. 5). The errors in palaeolatitudes of the base-maps have been obtained by giving unit weight to each rock unit when calculating the statistics of the mean polar wander path for North America¹² rather than to each determination of the geomagnetic field, that is, to each sampling site, of which there

may be many in an individual study. The combined error estimates of Fig. 5 may therefore be too large. To use a procedure (that of giving unit weight to each rock unit) that probably overestimates errors seems more prudent than to use a method (that of giving sites unit weight) that probably underestimates them.

We suggest that the most probable time of motion of the Stikine terrain from its more southerly position was latest Cretaceous or early Tertiary. The apparent absence of post-Triassic palaeolatitudinal motion of the Alexander terrain relative to the craton¹⁵ indicates that the Stikine block became welded in behind the Alexander terrain at that time. The older limit for the time of this motion is set by the Axelgold intrusion whose age uncertainty has already been discussed. An important additional argument in favour of this late motion comes from considering the palaeomagnetic evidence from the Coast Plutonic Complex (CPC of Fig. 1). The Coast Plutonic Complex is a belt of Cretaceous and early Tertiary intrusions and metamorphic bodies that has not been disrupted by later deformation. The Stikine terrain is attached to the complex because it is intruded by them. Magnetic inclinations observed in Cretaceous intrusions within the Coast Plutonic Complex (such as the Howe Sound pluton²⁷) are systematically too shallow, indicating that they were emplaced 10° or more south of their present position (see ref. 17, Fig. 3). Such data suffer from the absence of accurate corrections for possible post-emplacement tilt, but the effect is systematic over 400 km, and indicate a latest Cretaceous or earlier displacement consistent with our results.

A displacement of this magnitude, so late in the history of the Cordillera, and occurring so far to the east, raises the question of where the differential movement with stable North America is located. There is no geological evidence of such a late ocean closure between the localities sampled and the craton. The Cache Creek Group is thought to be ancestral Pacific Ocean floor which was thrust westward over the eastern margin of the Stikine terrain in post-Oxfordian time. Three major faults lie east of the localities sampled, and on two of them dextral strike-slip movement has been reported. The Pinchi Fault (Fig. 2) may have been a transform fault in the late Triassic, and right-lateral slip has probably occurred along it in the late Cretaceous and/or early Tertiary^{28,29}. The Swannell Fault, situated to the east of the Pinchi, juxtaposes metamorphosed Proterozoic strata against later Palaeozoic and early Mesozoic volcanogenic rocks, and may mark the ancient cratonic margin at this point. No evidence of lateral slip has been reported from it. The northern Rocky Mountain Trench is the site of the third major fault (Fig. 2). Almost certainly right-lateral slip has occurred on it, but the amount

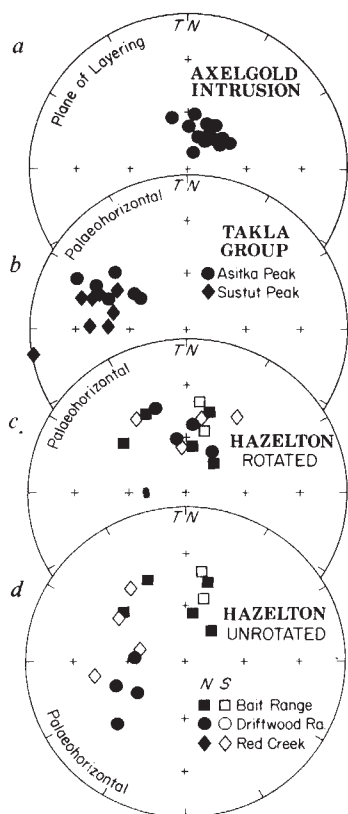
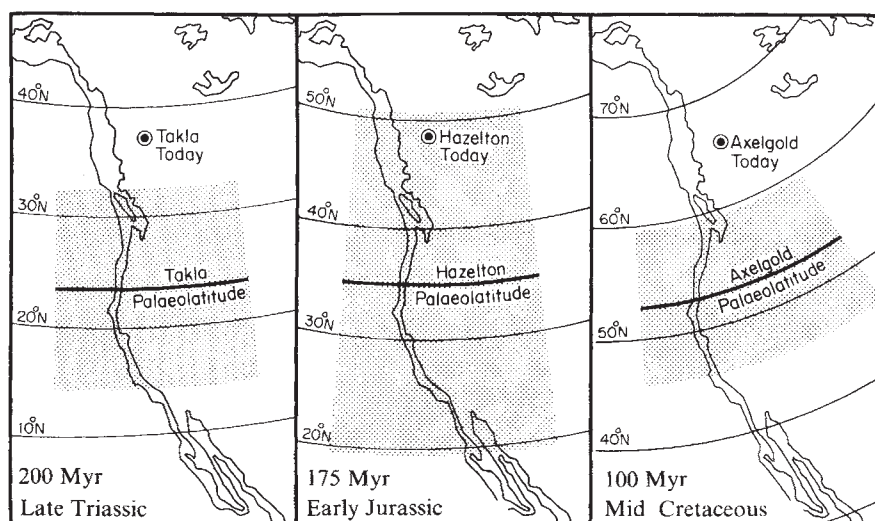


Fig. 4 Mean directions of characteristic magnetization at each collecting site. Directions are plotted on the lower hemisphere and the perimeter is the palaeohorizontal. The magnetization have normal polarity (north-seeking directions plotted) in the Axelgold and Takla rocks. In the Hazelton both polarities occur and the reversed directions (south-seeking directions, S) are shown by open symbols. In *d* Hazelton results are shown unrotated, correction being made for geological dip only. In *c* the directions are 'corrected' for rotations about local vertical axes (see text).

Fig. 5 Palaeolatitudes of the Stikine terrain compared with those calculated from cratonic North America. The reference palaeopoles¹² used for North America are for 200 Myr 68°N, 093°E, for 175 Myr 78°N, 110°E, for 100 Myr 69°N, 188°E. The error zones ($P=0.05$) combine the errors in both base maps and palaeolatitude observations from the Stikine terrain. Error estimates are probably too large (see text).



is unknown. It probably extends northwards into the Yukon Territory as the Tintina Fault (Fig. 1). Between 350 and 500 km of right-lateral movements since the mid-Cretaceous has been reported from the Tintina Fault^{30,31}, and the displacement may, in part, continue down the northern Rocky Mountain Trench and, in part, diverge westwards down such faults as the Pinchi and the Frazer lineaments (F, Fig. 1). The motions that can be documented geologically are much less than the 1,300 km derived from our evidence. For reasons discussed elsewhere³² Wrangellia was very probably attached to the Stikine terrain (Fig. 1) and rode northwards with it in the late Mesozoic or early Tertiary³².

Our interpretation conflicts with that of Symons³³ who argues that the area "has not been tectonically rotated or translated relative to stable North America". His arguments are based on his palaeomagnetic results from the Topley intrusions of the Stikine terrain. These intrusions are Jurassic in age, probably somewhat younger than the Hazelton Group. The Topley region has been intruded and partially covered by early Tertiary igneous rocks which have yielded K–Ar ages of between 50 and 63 Myr, and there was extensive hydrothermal activity and mineralization of economic importance at this time. We believe that Symons has misinterpreted his data, which are in fact consistent with our own. Symons' results (ref. 33, Fig. 3 and Table 1) fall into two distinct groups (*A* and *B*) which we believe have two separate origins. *A* magnetizations occur at eight sites in four of the Topley intrusions. They have a high mean stability index ($SI=20$) (ref. 33), normal and reversed polarity, and a mean inclination of 47°, corresponding to a palaeolatitude of 30° indistinguishable from that we find for the Hazelton (Fig. 5). The declinations are scattered and range from 294 to 348°, indicating variable anticlockwise rotations among the four Topley intrusions similar to the rotations we find for the Hazelton. The *A* magnetizations are, therefore, very similar to the magnetization of the Hazelton Group, and we suggest that they are primary and represent the Jurassic geomagnetic field. *B* magnetizations occur at seven sites in three other Topley intrusions³³. Their polarity is always normal, they have a low mean stability index ($SI=4$), and are well grouped with much steeper inclinations than *A* ($D=336^\circ$, $I=74^\circ$; $N=7$; $k=96$; $\alpha_{95}=6^\circ$). The corresponding pole, 75°N, 176°W, is statistically identical with the mean pole for North America at 60 Myr (77°N, 173°W; $\alpha_{95}=7^\circ$) (ref. 12). Hence, the less stable *B* magnetization was apparently acquired during the early Tertiary when igneous and hydrothermal activity was widespread in the Topley area. Symons averages all his results together and obtains a mean inclination of 63°, corresponding to a palaeolatitude of 40° $\pm 11^\circ$, which is not significantly different from that expected

from cratonic North America. We suggest that his average result has no tectonic significance because it is derived from combining a Jurassic magnetization (*A*) with a Tertiary magnetization (*B*).

Although the continuity of rock units indicates that the Stikine terrain has behaved as a coherent entity, it has not behaved as a rigid plate, because the declinations of the magnetizations directions are highly variable (Table 1). On its eastern margin the Axelgold Intrusion has rotated 63° clockwise relative to cratonic North America (rotations were calculated using equation (9.2) ref. 18). Clockwise rotations are very common in the western Cordillera, and have been ascribed to motions of small blocks activated by right-lateral shear of the Pacific and North American plates—the so-called 'ball-bearing' hypothesis³⁴. However, our palaeomagnetic declinations show large, variable and commonly anticlockwise rotations; 43° and 62° for the two Takla localities, 41° and 103° for two of the Hazelton localities and 14° clockwise for the third. Of course, the observed western declinations could be caused by much larger clockwise rotation (360° less the above angles) and so bring their sense of rotation into line with those generally found in the western Cordillera. However, the simplest explanation is that different parts of the Stikine terrain have undergone variable amounts of predominantly anticlockwise rotation relative to cratonic North America. This is inconsistent with the simple ball-bearing hypothesis³⁴. The Stikine terrain at least in its eastern part seems to be an assemblage of slices of crust of continental thickness with areal dimensions comparable to or less than the thickness of the lithosphere, and which are separated by faults with high strains (Fig. 2). Within each slice the strata are only gently tilted or folded²⁴.

We thank Martin Irving for help with rock collecting, Jean Hastie for magnetic measurements, Ray Yole for discussions, and Jean Roy for helpful criticisms.

Received 4 December 1979; accepted 18 March 1980.

- Hamilton, W. H. *Bull. geol. Soc. Am.* **80**, 2409–2430 (1969).
- Davis, G. A., Monger, J. W. H. & Burchfiel, B. C. in *Mesozoic Paleogeography Western United States* (eds Howell, D. G. & McDougall, K. A.) 1–32 (Society Economic Paleogeography and Mineralogy, Pacific Section, 1978).
- Monger, J. W. H. & Price, R. A. *Can. J. Earth Sci.* **16**, 770–791 (1979).
- Armstrong, R. L. *Geol. Ass. Can. Cordilleran Sect. Prog. Abstr.* **7** (1979).
- Jones, D. L., Silberling, N. J. & Hillhouse, J. W. *Can. J. Earth Sci.* **14**, 2565–2577 (1977).
- Monger, J. W. H. *Can. J. Earth Sci.* **14**, 1832–1859 (1977).
- Tozer, E. T. *Geol. Surv. Can. Econ. Geol. Rep.* **1**, 633–640 (1970).
- Monger, J. W. H. & Ross, C. A. *Can. J. Earth Sci.* **8**, 259–278 (1971).
- Irving, E. & Yole, R. W. *Earth Phys. Br. Publ. Energy, Mines and Resources, Can.* **42**, 87–95 (1972).
- Packer, D. R. & Stone, D. B. *Can. J. Earth Sci.* **11**, 976–997 (1974).
- Hillhouse, J. W. *Can. J. Earth Sci.* **14**, 2578–2592 (1977).
- Irving, E. *Can. J. Earth Sci.* **14**, 669–694 (1979).
- Muller, J. E. *Can. J. Earth Sci.* **14**, 2062–2085 (1977).

14. Berg, H. C., Jones, D. L. & Richter, D. H. *U.S. Geol. Surv. Prof. Pap.* **800D**, 1-24 (1972).
15. Hillhouse, J. W. *Geol. Soc. Am. Prog. Abstr. Meet.* **444** (1979).
16. Beck, M. E. & Nason, L. *Nature* **235**, 11 (1972).
17. Symons, D. T. A. *Can. J. Earth Sci.* **8**, 1388-1396 (1971).
18. Irving, E. *Paleomagnetism* 1-399 (Wiley, New York, 1964).
19. Simpson, R. W. & Cox, A. *Geology* **5**, 585-589 (1977).
20. Kamerling, M. & Luyendyk, B. P. *Bull. geol. Soc. Am.* **90**, 331-337 (1979).
21. Monger, J. W. H., Richards, T. A. & Paterson, I. A. *Can. J. Earth Sci.* **15**, 823-830 (1978).
22. Monger, J. W. H. *Geol. Surv. Can. Pap.* 76-29 (1977).
23. Wilson, R. L., Haggerty, S. E. & Watkins, N. D. *Geophys. J. R. Astr. Soc.* **16**, 79-5 (1968).
24. Tipper, H. W. & Richards, T. A. *Bull. geol. Surv. Can.* **270** (1976).
25. Irvine, T. N. *Geol. Surv. Can. Pap.* **75-1B**, 81-88 (1975).
26. Larson, R. & Pitman, W. J. *geophys. Res.* **82**, 3645-3662 (1972).
27. Symons, D. T. A. *Nature* **241**, 59-61 (1973).
28. Paterson, I. A. *Geol. Surv. Can. Paper* **74-1B**, 29-42 (1974).
29. Paterson, I. A. *Can. J. Earth Sci.* **14**, 1324-1342 (1977).
30. Roddick, J. A. *J. Geol.* **75**, 23-33 (1967).
31. Gabrielse, H. & Dodds, C. J. *Geol. Ass. Can. Prog. Abstr.* **2**, 19 (1977).
32. Irving, E., Monger, J. W. & Yole, R. W. *Geol. Ass. Can. Spec. Publ.* (in the press).
33. Symons, D. T. A. *Can. J. Earth Sci.* **10**, 1099-1108 (1973).
34. Beck, M. E. *Am. J. Sci.* **276**, 694-712 (1976).
35. Fisher, R. A. *Proc. R. Soc. A* **217**, 295-305 (1953).

Different species of messenger RNA encode receptor and secretory IgM μ chains differing at their carboxy termini

Paul A. Singer*, Helen H. Singer* & Alan R. Williamson

Department of Biochemistry, University of Glasgow, Glasgow, G12 8QQ, UK

Biosynthetic studies in the presence of an inhibitor of glycosylation indicate that individual human lymphoma-derived cell lines can synthesize both membrane receptor and presumptive secretory forms of IgM μ chains. The receptor form has a larger polypeptide chain than the secretory form and possesses a different C-terminus, but similar N-terminus, consistent with the presence of a C-terminal hydrophobic 'tail' for integral membrane binding. Messenger RNA isolated from these cells directs the synthesis of both forms of μ chain in a wheat germ translation system, indicating the presence of independent mRNAs for each form. It is proposed that the synthetic pathways for receptor and secretory IgM diverge at the post-transcriptional level, possibly by differential RNA splicing to give mRNA molecules with or without a translatable 'tail' segment.

RECEPTOR immunoglobulin molecules on the surface of B lymphocytes behave in many respects like typical integral membrane proteins. By contrast, serum immunoglobulins, which function as effector antibodies, are hydrophilic globular proteins. It is reasonable to postulate that this dichotomy of function of immunoglobulins is dictated by a duality of structure. For instance, the immunoglobulin heavy chain may be expected to possess restricted regions of hydrophobicity which would function in membrane interaction. However, serum immunoglobulins do not seem to possess such hydrophobic regions. This observation has led previous workers to postulate the presence of a hydrophobic peptide on the receptor form of immunoglobulin, which is absent from the otherwise similar secretory form^{1,2}. Alternative explanations have also considered the possibility that differences in glycosylation of the receptor form could cause conformational changes which would favour membrane binding².

There has been a steady accumulation of indirect evidence as to the manner of membrane binding by receptor immunoglobulin. The μ chain from radioiodinated splenocyte receptor IgM has been shown to be of higher molecular weight than its secreted counterpart by SDS-polyacrylamide gel electrophoresis (SDS-PAGE)³. We have made a similar observation for receptor IgM and IgG μ and γ chains synthesized in several human lymphoid cell lines⁴, and a similar phenomenon has been reported for receptor IgM of a mouse lymphoid cell line⁵. Buoyant density and solubility measurements in the presence of non-ionic detergents have shown a greater hydrophobicity for receptor IgM, as opposed to secretory IgM^{6,7}. In spite of these findings, the basis for the observed differences in receptor and secretory immunoglobulins remained unclear. The apparent molecular weight differences could be due either to variation in the

extent of glycosylation or to a difference in polypeptide chain lengths. More recently, carboxypeptidase digestion and peptide mapping techniques have been used in several laboratories^{5,8,9} to analyse membrane associated μ chains for a modified C-terminus. These studies have so far yielded conflicting results, which, in some experiments, may be due to contamination by the secretory μ -chain species.

A separate but related question to the structure of receptor immunoglobulin concerns the biosynthetic pathway which distinguishes it from secretory immunoglobulin. Previous models^{10,11} have proposed that divergence of the pathways occurs just before secretion, based on the finding that intracellular transport of both forms is very similar^{12,13}. However, investigations of the events preceding synthesis of the polypeptide chains have not been previously reported.

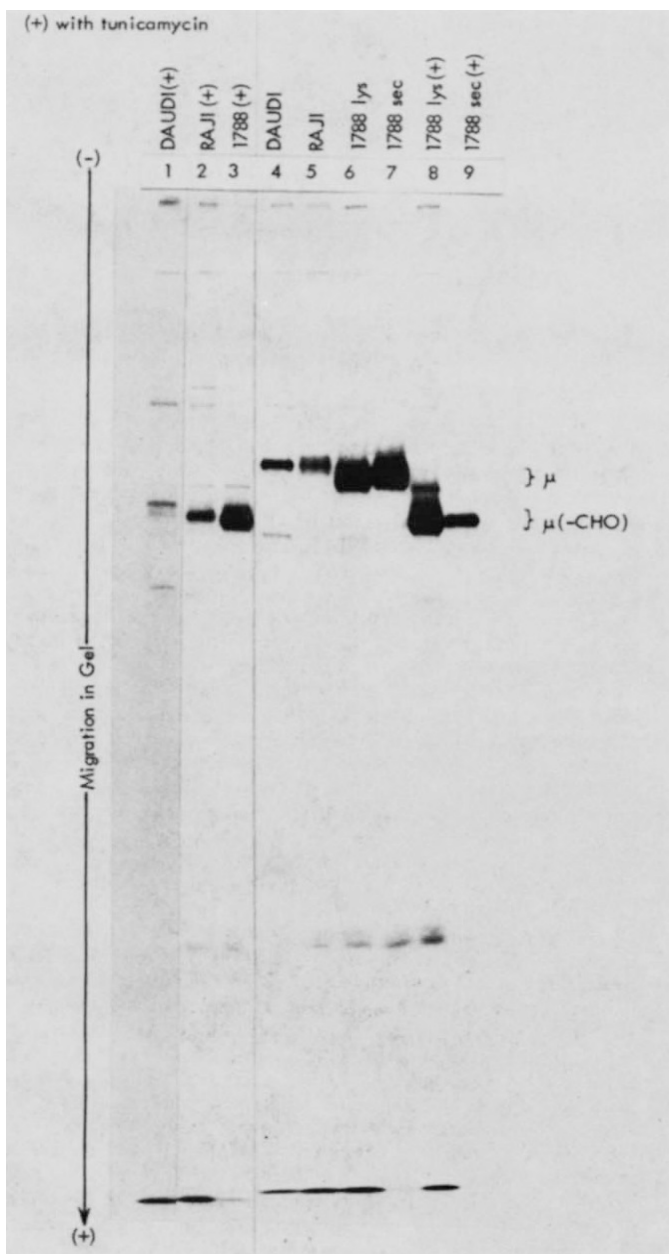
In this article we report on the structure and biosynthesis of receptor IgM chains in human lymphoid cells representing various stages of B-lymphocyte differentiation and immunoglobulin expression. We have used the drug tunicamycin as an inhibitor of glycosylation. Size comparisons of non-glycosylated μ chains indicate that the membrane form has a larger polypeptide chain than the secretory form. The larger membrane form of μ chain was isolated and shown by carboxypeptidase digestion to have a modified C-terminus (but a normal N-terminal sequence) relative to secretory μ chain, consistent with the presence on receptor μ chain of a C-terminal peptide, possibly hydrophobic in nature, capable of interaction with the plasma membrane. Unexpectedly, two forms of μ chain were detected intracellularly in Daudi and Raji nonsecretory cells, one resembling the membrane form and one resembling the secretory form with the former predominating in Daudi cells and the latter form predominating in Raji cells. This same pattern of μ -chain expression was found by cell-free translation of Daudi and Raji mRNA, indicating that the two forms of μ chain can be

*Present address: Max-Planck-Institut für Biologie, Corrensstrasse 42, 7400 Tübingen 1, FRG.

produced in the same cell by simultaneous translation of different mRNA molecules.

Cell lines with qualitatively and quantitatively different immunoglobulin biosynthetic capacities were chosen for examination (Table 1). Raji and Daudi cells each devote about 0.5% of their total protein synthesis to immunoglobulin production. Daudi cells deposit most of this immunoglobulin on the plasma membrane. Raji cells, by contrast, deposit very little of the immunoglobulin on the surface in a form detectable either by fluorescent antibody probes or by surface radioiodination. This comparison suggests that Raji more closely resembles a pre-B lymphocyte, while Daudi appears to be an analogue of a more mature B lymphocyte. RPMI 1788 is a high-rate IgM secreting lymphoblastoid cell line.

Our findings on the intracellular synthesis of membrane and secretory type μ chains reveal several unexpected features in the intracellular processing of μ chains by Daudi and Raji cells (Table 1). These features are discussed further below, and include the expression in characteristic ratios of two forms of μ chain by cell-free translation of isolated mRNA and in tunicamycin-treated cells, apparent lag in processing of leader sequences as revealed in the presence of tunicamycin, and probable unequal glycosylation of the two μ -chain forms.



Analysis of the amino- and carboxy-terminal sequences of surface μ chains

The intracellular forms of μ chain found in nonsecretory and secretory cells can be distinguished from each other and isolated by immunoprecipitation and SDS-PAGE. Analytical characterization of the intracellular μ chains of Daudi, Raji and RPMI 1788 is shown in Fig. 1. The intracellular form of μ chain radiolabelled during a 1-h exposure of Daudi (lane 4) or Raji (lane 5) cells to ^{35}S -labelled methionine is of similar apparent molecular weight and is distinct from the lower molecular weight intracellular μ chain of the high rate secreting cell RPMI 1788 (lane 6).

The higher molecular weight of the surface μ chain could be due to either extra carbohydrate or an extra peptide at either amino-terminal or carboxy-terminal ends of the secreted μ chain. The hypothesis that surface μ chain contains an extra peptide was tested by analysis of the amino- and carboxy-terminal amino acid sequences of biosynthetically labelled μ chains. The partial amino-terminal sequences of Daudi and Raji μ chains (Fig. 2) show that each has a pattern of valine and leucine residues corresponding to the prototype $\text{V}_{\mu\text{III}}$ subgroup sequence. The period of labelling with radioactive valine and leucine was shorter than the turnover time for membrane immunoglobulin. Consequently these sequences establish that neither Raji nor Daudi μ chains retain the hydrophobic amino-terminal precursor peptide sequence, with which H chain synthesis is initiated, long enough for it to serve as a surface membrane insertion site. The μ chain of RPMI 1788 was subjected to similar analysis but proved to have a blocked amino terminus. In the absence of other sequence information, this blocked amino terminus is consistent with processing of the precursor peptide as expected for a secretory μ chain.

For comparative carboxy-terminal analysis of surface and secreted μ chains the enzyme carboxypeptidase-A was used. The carboxy-terminal residue to secretory μ chain is tyrosine¹⁴, and carboxypeptidase-A will rapidly and completely remove a C-terminal tyrosine residue¹⁵. If surface μ chains possess an extra carboxy-terminal peptide they would not be expected to have a terminal tyrosine residue. This

Fig. 1 SDS-PAGE of biosynthetically labelled μ chains from tunicamycin treated and untreated cells. Log phase cultures were incubated with tunicamycin ($0.5 \mu\text{g ml}^{-1}$) for 4 h. Parallel untreated cultures were also maintained. 1×10^6 cells were collected for a 1-h pulse labelling in 0.2 ml medium containing no exogenous methionine and supplemented with $100 \mu\text{Ci } ^{35}\text{S}$ -methionine (500 Ci mmol^{-1}). Complete RPMI 1640 medium (0.8 ml) supplemented with 15% fetal calf serum (growth medium) was added and incubation continued for 2 h. For the treated cultures, tunicamycin ($0.5 \mu\text{g ml}^{-1}$) was present throughout the labelling and chase periods. Cells were separated from culture medium and lysed by the addition of 0.2 ml ice cold lysis buffer (0.15 M NaCl , 1% Nonidet P-40, 1 mM phenylmethylsulphonylfluoride, pepstatin A ($0.7 \mu\text{g ml}^{-1}$), 25 mM Tris-HCl pH 8.2). Cell lysates were successively centrifuged to remove nuclei ($1,000\text{g}$, 10 min) and cellular debris ($30,000\text{g}$, 30 min) and labelled IgM was immunoprecipitated by addition of $10 \mu\text{g}$ affinity purified rabbit anti-human IgM antibody (15 min at 20°C) and 25 μl of 10% solution of formalin-fixed, protein A-bearing *Staphylococcus aureus* organisms (Cowan I strain)³⁵. The immunoprecipitates were washed three times with 0.5-ml aliquots of lysis buffer, heated to 100°C for 2 min in SDS-PAGE sample buffer (2% SDS, 0.1 M dithiothreitol, 10% glycerol, 0.01 mg ml^{-1} bromophenol blue dye, 0.065 M Tris-HCl pH 6.8), and electrophoresed in a 10% polyacrylamide slab gel according to the method of Laemmli³⁶. Radioactive components were visualized by fluorography³⁷ (exposures of 2–24 h). The positions of glycosylated and non-glycosylated μ chains are indicated in the margin. Lanes 1, 2, 3, 8 and 9, tunicamycin treated; lanes 4, 5, 6 and 7, untreated cells.

Raji	-V-LV-----LV-----L-L
Daudi	-V-LV-----LV-----L-L
V _H III	EVQLVESGGGLVQPGGSLRL

Fig. 2 Partial amino-terminal amino acid sequences of μ chains. The μ chains of Raji and Daudi cell IgM were biosynthetically labelled with ^3H -leucine (55 Ci mmol^{-1}) and ^3H -valine (36 Ci mmol^{-1}) and isolated by immune precipitation and SDS-PAGE as described in legends to Figs 1 and 3. Amino acid sequences were determined by automated Edman degradation performed on a Beckman 890C sequencer. Repetitive yield at each cycle determined by internal standards was 88–92%. Separate sequencer runs were performed for each sample labelled with each amino acid. The V_HIII prototype sequence is from ref. 38.

approach has already been used to investigate the carboxy terminus of μ chains from the human cell line Bri-8 and the mouse myeloma tumour McPc 1748 (ref. 8), as well as the mouse B cell line 38C-13 (ref. 5). In all three cases there was significant release of radioactive tyrosine from μ chain by carboxypeptidase-A treatment. However, in those studies there was no convincing demonstration that the μ chains investigated were of the surface form. Also there was a lack of kinetic evidence to support the proposal that the tyrosine residues released were indeed carboxy-terminal. The Bri-8 cell line studied in this laboratory is a high-rate IgM secreting line. Recently carboxypeptidase-A digestion of Daudi μ chain has been reported to release fractional molar amounts of several amino acids, including tyrosine, but kinetics suggested that valine is located closer to the carboxy terminus than tyrosine⁹. For the present studies Daudi, Raji, RPMI 1788 and Bri-8 cells were labelled for 1 h with ^3H -tyrosine and the intracellular μ chains and L chains isolated by immunoprecipitation and SDS-PAGE. Patterns similar to those shown in Fig. 1 (lanes 4–6) were obtained and individual μ chains and light chains were eluted from the appropriate regions of the gel. The kinetics of carboxypeptidase-A catalysed release of ^3H -tyrosine from these isolated chains is shown in Fig. 3. For the μ chains of secreted IgM of RPMI 1788 and Bri-8 approximately 1 mol of tyrosine is released per chain during 2 min of incubation. This finding is consistent with the kinetics of release of carboxy terminal tyrosine from aldolase¹⁶. Carboxypeptidase analysis of the ^3H -tyrosine labelled light chains of RPMI 1788, Bri-8 and Raji shows no release of tyrosine. This is consistent with known sequences and established the low background in this analysis. The intracellular μ chains of characteristic surface size isolated from Daudi and Raji cells showed a partial release of tyrosine in the carboxypeptidase-A analysis (Fig. 3). Tyrosine was released from Daudi and Raji μ chains with kinetics indicating a carboxy-terminal location, but with the yield in both cases significantly lower than expected. The yields of tyrosine released correspond to approximately 0.35 mol for Daudi and 0.7 mol for Raji μ chain (based on an estimate of 20 tyrosine residues per chain). The simplest interpretation of these results is that Daudi and Raji cells contain, in different proportions, two different forms of μ chain, only one of which has a carboxy-terminal tyrosine residue.

Analysis of non-glycosylated μ chains

Human μ chains are glycosylated at five asparagine residues and the carbohydrate side chains constitute approximately 15% of the molecular weight of human IgM^{17–19}. To test the hypothesis that the difference in molecular weight between

surface and secreted μ chains could be due to a difference in the extent of glycosylation, the drug tunicamycin was used to inhibit glycosylation during μ chain biosynthesis. Tunicamycin has been shown to act by inhibiting the formation of *N*-acetylglucosamine lipid intermediates²⁰, which serve as donors for the synthesis of the core regions of asparagine-linked oligosaccharides²¹. Biosynthesis of completely non-glycosylated μ chains was effected by preincubating the lymphoid cells with tunicamycin in limiting conditions²². SDS-PAGE analyses of the non-glycosylated μ chains are shown in Fig. 1 and in more detail in Fig. 4. The intracellular form of secretory μ chain from tunicamycin treated RPMI 1788 cells is a single band (lane 9, Fig. 4) with a considerably lower apparent molecular weight than the glycosylated intracellular μ chain of RPMI 1788 (lane 3) or the glycosylated

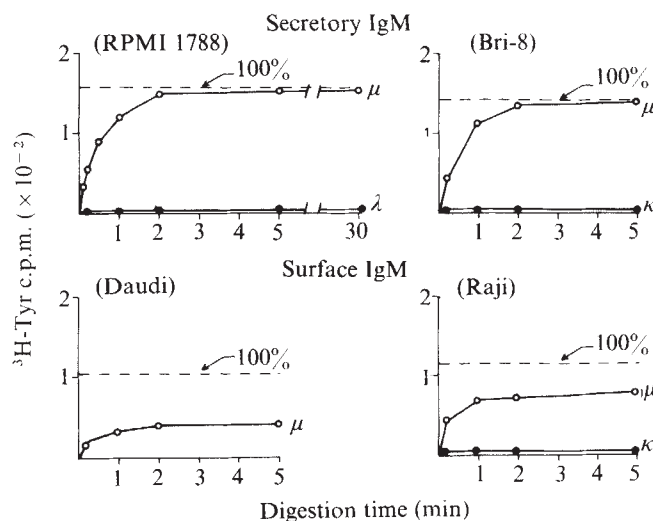


Fig. 3 Carboxy-terminal tyrosine released from intracellular μ chains of the pre-membrane and pre-secretory types. Log phase cells (5×10^6 per ml) were labelled for 1 h in either 0.2 ml (Bri-8, RPMI 1788) for 0.5 ml (Daudi, Raji) RPMI 1640 lacking exogenous tyrosine and supplemented with ^3H -tyrosine ($500\text{ } \mu\text{Ci ml}^{-1}$; 40 Ci mmol^{-1}). Cell lysis, anti- μ immunoprecipitations, SDS-PAGE and fluorography were done as described for Fig. 1. The fluorography of the gel (not shown) revealed μ and light chain bands similar to those shown in Fig. 1, lanes 4, 5 and 6 (Bri-8 μ -chain mobility was indistinguishable from that of RPMI 1788). Bands were located by overlaying the dried gel with the fluorographic negative. Gel pieces containing the bands were sliced out and the material eluted overnight at 60°C in 0.5 ml elution buffer (2% SDS, 20 mM dithiothreitol, 0.25 M Tris-HCl, pH 8.2). After removal of the gel slices, 0.5 mg bovine serum albumin (as carrier) and freshly recrystallized iodoacetamide (final concentration 50 mM) were added and the samples incubated in the dark at 20°C for 30 min. The proteins were precipitated with acetone (five volumes), collected by centrifugation, and washed twice with acetone:water (5:1 v/v). The air-dried precipitates were dissolved by heating at 100°C for 2 min in 0.9 ml digestion buffer (0.2 M LiCl, 50 mM NaHCO_3). Insoluble material (less than 10% of the radioactivity) was removed by centrifugation, and the samples adjusted to pH 8.0 with 0.5 M NaOH. Crystals of carboxypeptidase A (Worthington Biochemicals) were washed with water and dissolved in 10% LiCl at 0.2 mg ml^{-1} ($A_{280}/2.3 = \text{mg ml}^{-1}$). Digestions (enzyme:substrate = 1:100) were initiated by addition at 20°C of 10 μl enzyme solution with rapid mixing. Aliquots (150 μl) of reaction mixture were removed at the times indicated, immediately added to 50 μl ice cold 40% trichloroacetic acid, for determination of released ^3H -tyrosine (acid-soluble) fraction. The hypothetical 100% release values are based on an estimated 20 tyrosine residues per μ chain (or 18 for RPMI 1788).

Table 1 Summary of μ -chain expression in the cell lines investigated

	IgM expression*		Level†	μ Chains detected intracellularly‡		μ chains detected by cell-free translation
	Membrane	Secretory		Glycosylated	Non-glycosylated	
Daudi	+++	—	0.5%	{ MEMBRANE } secretory	PREMEMBRANE presecretory	PREMEMBRANE presecretory
Raji	+	—	0.5%	{ membrane } SECRETORY	premembrane PRESECRETORY	premembrane PRESECRETORY
RPMI 1788	±	+++	5–10%	SECRETORY	PROCESSED SECRETORY	PRESECRETORY

CAPITALS designate the predominant or only form detected.

*Qualitative expression as determined by using fluorescent antibody, surface radioiodination and biosynthesis techniques.

†Incorporation of ^{35}S -methionine into anti- μ precipitable relative to trichloroacetic acid precipitable cellular proteins during a 60-min labelling.

‡Species labelled with ^{35}S -methionine during a 15–60-min pulse, resolved by SDS-PAGE and/or characterized by C-terminal analysis. Terminology as follows: membrane—membrane μ chain lacking C-terminal tyrosine; secretory—secretory μ chain with C-terminal tyrosine; presecretory and premembrane—respectively μ -chain forms with presumed N-terminal leader sequence; processed secretory—indicates cleavage of N-terminal leader sequence. The presence or absence of the N-terminal leader sequence (that is, yielding pre or processed μ chain respectively) can be assessed by comparison of the mobilities on SDS-PAGE of non-glycosylated μ chains with the products of cell-free translation of μ -chain mRNA.

§These are detected by SDS-PAGE as one band of membrane μ -chain size, however C-terminal analysis of glycosylated chains is consistent with the indicated composition. The processing of glycosylated chains cannot be resolved by SDS-PAGE analysis but amino-terminal sequence analysis (Fig. 2) shows that processing is complete within 6 h after synthesis.

extracellular μ chain (lane 7, Fig. 1). The non-glycosylated μ chains of Raji and Daudi intracellular IgM are resolved into two forms (lanes 7, 8, Fig. 4) differing from one another in their apparent molecular weights. The lower molecular weight μ chains of Daudi and Raji are similar in mobility to each other and have a perceptibly lower mobility than the non-glycosylated μ chain of RPMI 1788 cells (lane 9). The major difference between Daudi and Raji profiles is in the intensity of the labelling of the two forms of μ chain, with the higher molecular weight form of μ chain predominating in Daudi, and the lower molecular weight form predominating in Raji. The existence of two forms of μ chains in Daudi and Raji cells confirms the prediction made on the basis of the carboxy-terminal analysis as discussed above. Comparison of the relative proportions of the two forms of μ chain with the differential release of tyrosine from Daudi and Raji μ chains is consistent with the interpretation that the higher molecular weight form of both Daudi and Raji μ chain lacks carboxy-terminal tyrosine and the lower molecular weight form has a carboxy-terminal tyrosine residue (thus being analogous to the secreted μ chain of similar apparent molecular weight). On this hypothesis the higher molecular weight form of μ chain in Raji and Daudi cells is the precursor of surface μ chain, and the molecular weight difference relative to the secretory μ chain is due to an extra carboxy-terminal peptide attached to the secretory μ -chain sequence and preventing carboxypeptidase catalysed release of the carboxy-terminal tyrosine. Evidence consistent with this hypothesis was obtained by separation of the two forms of non-glycosylated Daudi μ chain biosynthetically labelled with tyrosine and subsequent carboxypeptidase digestion. The higher molecular weight form of μ chain has essentially no carboxypeptidase-releasable tyrosine while the lower molecular weight form of μ chain yields a mole of tyrosine and kinetics consistent with tyrosine being the carboxy-terminal residue (Fig. 5). A similar analysis of the two Raji μ chains (not shown) yielded similar results.

A comparison of the SDS-PAGE profiles of Daudi and Raji non-glycosylated μ chains with those of glycosylated μ chains (compare lanes 7 and 8, Fig. 4 with lanes 4 and 5, Fig. 1) shows that glycosylation abolishes the mobility differences observed between the membrane and secretory forms. This suggests that during a 1 h incorporation of label the secretory-type μ chains from Daudi and Raji cells have more

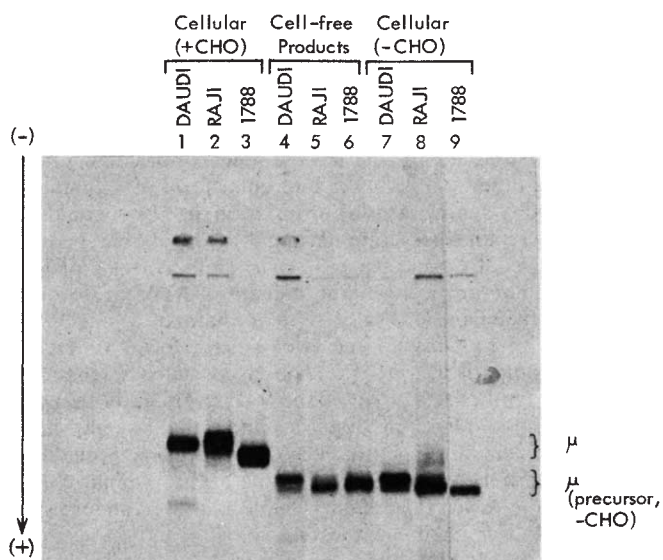


Fig. 4 SDS-PAGE comparing membrane and secretory type μ chains synthesized *in vivo* in glycosylated (+CHO) or non-glycosylated (–CHO) forms and in a wheat germ cell-free system. Cells were labelled by incorporation of ^{35}S -methionine in the presence (for 60 min) or absence (for 15 min) of tunicamycin ($0.5\mu\text{g ml}^{-1}$) as described in the legend to Fig. 1. Wheat germ extract was prepared and assays run essentially as described previously²⁵. Assays ($50\mu\text{l}$) contained $25\mu\text{Ci}$ ^{35}S -methionine (800Ci mmol^{-1}) and poly(A)⁺ RNA ($1\mu\text{g}$) prepared from each cell line by phenol extraction and oligo(dT)-cellulose chromatography of isolated polyribosomes. After incubation, assays were diluted fivefold with 3D-TKM buffer (0.1M KCl , 5mM MgCl_2 , 1% Triton X-100, 1% sodium deoxycholate, 0.5% SDS, $0.1\text{M Tris-HCl pH } 8.0$). Anti- μ immunoprecipitations were carried out on cell lysates and diluted wheat-germ assays, and the immunoprecipitates analysed by SDS-PAGE on reducing 10% polyacrylamide slab gels as described in the legend to Fig. 1. The approximate mobility ranges of glycosylated, nonglycosylated and cell-free synthesized μ chains are indicated in the margin.

carbohydrate attached relative to the membrane μ chains. Mobility comparisons with the RPMI 1788 intracellular μ chain (lane 6, Fig. 1) support the idea that it is overglycosylation of the Daudi and Raji secretory-type μ chains, and not underglycosylation of the membrane forms, which is occurring. A more direct experimental approach will be required to resolve this point, however.

Cell-free biosynthesis of μ chains

The above results are consistent with, but do not prove, an independent biosynthetic origin for the two forms of μ chain in Daudi and Raji cells. A separate biosynthetic origin for these two μ chains was confirmed by their production in a cell-free protein synthesis system.

Messenger RNA capable of being efficiently translated in a wheat-germ cell-free system was isolated from Daudi, Raji and RPMI 1788 cells. Polyribosomes were prepared and polyribosomal RNA fractionated over oligo(dT)-cellulose. The poly(A) containing mRNA was translated in a wheat-germ cell-free system and the immunoglobulin chains produced were specifically immunoprecipitated and analysed by SDS-PAGE (Fig. 4). Daudi and Raji μ chains resolved into two bands (lanes 4 and 5).

These two forms of μ chains are similar in apparent molecular weight to the two non-glycosylated forms of μ chain synthesized *in vivo* in the presence of tunicamycin (compare lanes 4 and 5 with lanes 7 and 8, Fig. 4). It is also striking that the relative labelling intensities of the two forms of μ chain resulting from cell-free translation of Daudi and Raji messenger RNA resembles the labelling pattern seen *in vivo* for the non-glycosylated μ chains. The strong similarity of the patterns seen *in vitro* and *in vivo* supports the hypothesis that both Daudi and Raji cells contain two different μ -chain mRNA molecules, one coding for the higher molecular weight form and the other for the lower molecular form of μ chain.

The μ chains synthesized by translation *in vitro* would be expected to retain the amino-terminal hydrophobic peptide that is removed by processing *in vivo*. Data consistent with this phenomenon are seen for the μ chain of RPMI 1788. The cell-free product (lane 6, Fig. 4) is a μ chain of a single size intermediate in mobility to the two μ -chain forms of Daudi and Raji. Comparison with the non-glycosylated μ chain of RPMI 1788 synthesized *in vivo* (lane 9, Fig. 4) shows that the cell-free product differs in apparent molecular weight in a manner consistent with the presence on the cell-free μ chain of the amino-terminal precursor peptide. The similarity of mobility of the lower (and higher) molecular weight forms of Daudi and Raji μ chain synthesized *in vivo* and *in vitro* is consistent with the retention of the precursor peptide on the Daudi and Raji μ chains synthesized *in vivo* in the presence of tunicamycin. This would indicate that processing takes much longer than 1 h (the pulse-labelling time) but, in the absence of tunicamycin, processing is complete within 6 h (Fig. 2). Processing of an immunoglobulin κ chain can be retarded *in vivo* by culturing myeloma cells with protease inhibitors²³. The expected size difference between the precursor form of RPMI 1788 μ chain and the mature non-glycosylated form is approximately 2,000 daltons^{24,25}. This allows us to estimate the molecular weight difference between the two forms of Daudi and Raji μ chain as being approximately 3,000 daltons.

The size difference between the two forms of Daudi and Raji μ chain considered together with the demonstration of the absence of carboxy-terminal tyrosine in the higher molecular weight μ chain is consistent with the presence of an extra carboxy-terminal peptide of approximately 30 residues long relative to the lower molecular weight form of μ chain. The evidence of Williams *et al.*⁹ indicates that the carboxy-terminal peptide of Daudi surface μ chain is probably hydrophobic. It is postulated that the carboxy terminal hydrophobic peptide (tail peptide, T) serves to anchor the μ chain of surface IgM to

membrane from the point of synthesis to display of IgM on the plasma membrane. The T peptide provides a structural explanation for the difference between surface and secretory IgM and allows surface IgM to be an integral membrane protein resembling other integral membrane proteins in being attached to the membrane by a carboxy-terminal hydrophobic peptide.

Previously no direct evidence could be found for surface IgM being a transmembrane protein. The negative finding was based on the absence of iodination of μ chain on the inside of the plasma membrane²⁶ but those data can now be interpreted to mean that there is no tyrosine residue available in the T peptide. Two previous studies had claimed that surface μ chain has a carboxy-terminal tyrosine residue and therefore lacks a carboxy-terminal hydrophobic peptide for membrane attachment^{5,8}. Those results were probably due to analysis of a mixture of secretory and surface μ chains, although it is also possible that the tyrosine released by carboxypeptidase may have been located within the T peptide since kinetic evidence for carboxy-terminal location was not supplied. The present unexpected finding that Daudi and Raji cells synthesize two forms of μ chain means that a demonstration that a given cell line does not secrete immunoglobulin can no longer be accepted as evidence that the cell line will only synthesize the surface form of heavy chain. Both Raji and Daudi cells display IgM on the surface and neither secretes any IgM; however, both lines apparently synthesize an intracellular form of μ chain lacking the T peptide. For both cell types the quantitation is consistent with the hypothesis that only the μ chain with the extra T peptide is displayed on the cell surface but this remains to be proven by a chase experiment. We suggest that the secretory-sized form of μ chain synthesized in Daudi and Raji cells is metabolized intracellularly.

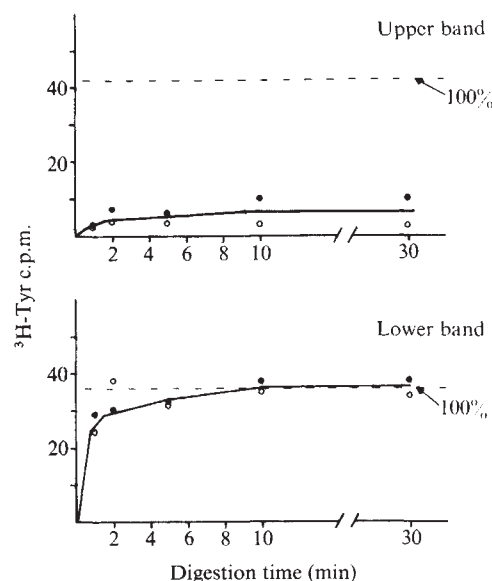


Fig. 5 Carboxy-terminal analysis of the larger (top) and smaller (bottom) forms of Daudi μ chains. Daudi cells were preincubated for 4 h in growth medium containing tunicamycin ($0.5 \mu\text{g ml}^{-1}$). Cell labelling with ^3H -tyrosine, lysis, immunoprecipitations, SDS-PAGE, fluorography and elution of radioactive bands were done as described for Fig. 3. The gel (not shown) revealed doublet bands of μ -chain mobility similar to those in Fig. 5, lane 7. For accurate separation of μ -chain species, the bands were sliced out one at a time, and their orientation relative to the main gel preserved. Fluorographs were taken after removal of each band and used to determine the accuracy of the slicing. Gel pieces were then trimmed as necessary to minimize cross-contamination. Carboxypeptidase digestions were performed and the results are plotted as described for Fig. 3. Reproducibility of the data is indicated by the results of duplicate determinations at each time point.

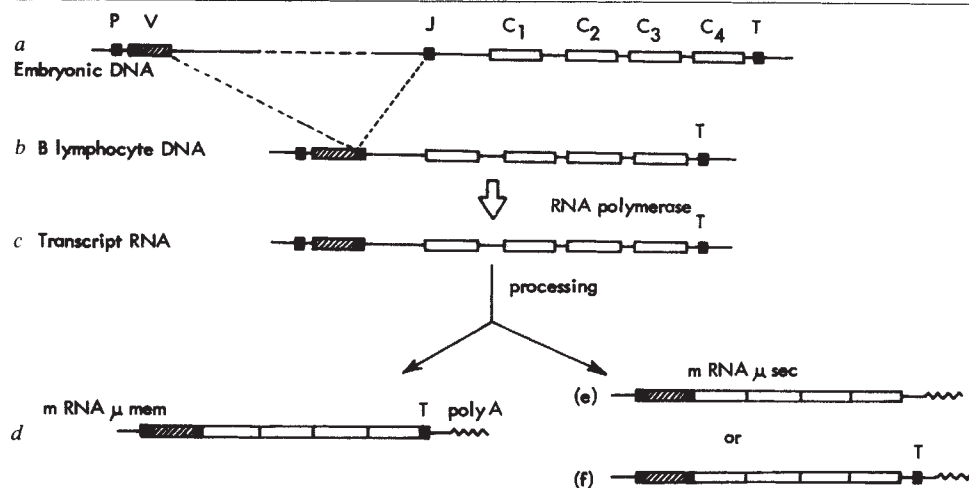


Fig. 6 Model for T-gene segment and the biogenesis of mRNA_{μmem} and mRNA_{μsec}. Depicted are the probable arrangements of heavy chain gene segments^{39,40} in embryonic DNA (a) and B-lymphocyte DNA (b), including the postulated T segment. A DNA rearrangement bringing the P + V_H segment into contiguity with the J_H + C_H gene segments, by analogy with V_λ rearrangement⁴¹, is shown. A similar rearrangement of the T-gene segment from embryonic to B-lymphocyte DNA is possible, but not necessary to the model. The possibility of a later

rearrangement of B-lymphocyte DNA, which would give rise to mRNA_{μmem} or mRNA_{μsec} is unlikely in view of the finding of simultaneous expression of both forms in individual cells. The model shows a single B-lymphocyte μ -chain transcription unit with the T segment in its embryonic relationship to the other C-region gene segments. Expression of the heavy chain transcription unit would yield a single common nuclear precursor RNA (c), which would then be processed by a splicing mechanism into mature mRNA_{μmem} (d, T segment joined to C_{H4}) and mRNA_{μsec} (e, T segment spliced out or f, T segment in non-translated RNA). Processing would presumably require specific recognition sequences and splicing enzymes, the activity of which would determine the relative production of membrane and secretory IgM. Each immunoglobulin heavy chain class could have a unique T segment, or a single T segment could be spliced into two or more heavy chain genes during RNA processing.

Our evidence indicates that the two forms of μ chain, with and without the T peptide, are produced by translation of different mRNA molecules. These two mRNA molecules could originate either from separate genes or from a single segmented gene by differential processing of a common nuclear RNA precursor. These models must be considered in the light of our knowledge of surface immunoglobulin and the switch from surface to secretory immunoglobulin. Surface receptor antibody on B lymphocytes can apparently be of any class or subclass of immunoglobulin²⁷ and indeed two classes of immunoglobulin, apparently associated with the same idiotype²⁸⁻³⁰ or antibody specificity³¹, can be found on the surface of a single cell. A suitable model must therefore allow the attachment of a T peptide to the constant region of any class or subclass of immunoglobulin and embrace the simultaneous expression of the surface form of two different classes of immunoglobulin. The segmented nature of an immunoglobulin H chain gene leads to the suggestion that the T peptide is encoded by a separate DNA segment, the T gene segment. Each C_H gene could have its own T segment or a single T segment could be shared by some or all of the C_H genes with rearrangement of the T segment to allow expression. The T segments would be selected to code for a peptide with a more consistently hydrophobic nature than that of the amino-terminal precursor peptide, but as with the precursor peptides, sequence diversity of the T segment might be expected. The postulated T-gene segment would function as part of the heavy chain transcription unit.

The simultaneous production of surface and secreted forms of the same immunoglobulin molecule by monoclonal cell lines in culture^{4,5} (or surface and intracellular forms as demonstrated in these studies) argues against rearrangement of DNA segments being required for the switch from surface to secreted immunoglobulin. Thus the model which best fits the present data is one in which the switch from surface to secreted H chain involves a change at the level of nuclear RNA processing (Fig. 6). It is consistent with this model that there is no evidence for different genetic markers on surface and secreted immunoglobulins of the same class. Changes in the processing of RNA transcripts have been evoked as a possible explanation for the switch from one immunoglobulin class to another during clonal expansion³². However, studies

on hybrid cells have so far failed to show evidence for an enzymatic mechanism involved in differential processing of μ and γ chain RNA³³. Recently, direct evidence against this mechanism for heavy chain class switching has been obtained from the analysis of large nuclear RNA with combinations of μ , γ and α gene probes³⁴.

Our present finding suggests that messenger RNA molecules coding for both μ_{mem} and μ_{sec} may function in either pre-B or B lymphocytes. It is also possible that the two types of μ_{sec} messenger RNA (Fig. 6e and f) could function at different stages of B-lymphocyte differentiation.

Since this manuscript was submitted two relevant reports have appeared. (1) The murine of γ_1 heavy chain gene has been completely sequenced by Honjo *et al.*⁴². This sequence shows a putative RNA splicing site near the 3' end of the C_{H3} coding sequence. (2) Four different sizes of μ chain mRNA have been resolved in a murine lymphoma line by Perry and Kelley⁴³. The translation products of these mRNA molecules have yet to be defined.

We thank Dr J. D. Capra for performing amino acid sequence analyses, and Drs Robert L. Hamill (Lilly Research) and W. F. J. Cuthbertson (Glaxo-Allenburys Research) for gifts of tunicamycin.

Note added in proof: Evidence supporting the mechanism illustrated in Fig. 6 has recently been obtained from sequences of a cloned DNA encoding the C_μ gene segments⁴⁴. Evidence for two mRNA molecules differing at their 3' ends comes from sequence of cloned cDNA corresponding to μ mRNA⁴⁵.

Received 16 August 1979; accepted 18 March 1980.

1. Singer, S. J. *Adv. Immun.* **19**, 1-66 (1974).
2. Vitetta, E. S. & Uhr, J. W. *Biochim. biophys. Acta* **415**, 253-271 (1975).
3. Melcher, U. & Uhr, J. W. *J. Immun.* **116**, 409-415 (1976).
4. Singer, P. A. & Williamson, A. R. *Eur. J. Immun.* (in the press).
5. Bergman, Y. & Haimovich, J. *Eur. J. Immun.* **8**, 876-880 (1978).
6. Melcher, U., Eidels, L. & Uhr, J. W. *Nature* **258**, 434-435 (1976).
7. Melcher, U. & Uhr, J. W. *Biochemistry* **16**, 145-152 (1977).
8. McIlhinney, R. A. J., Richardson, N. E. & Feinstein, A. *Nature* **272**, 555-557 (1977).
9. Williams, P. B., Kubo, R. T. & Grey, H. M. *J. Immun.* **121**, 2435-2439 (1978).
10. Uhr, J. W. *Cell Immun.* **1**, 228-244 (1970).
11. Uhr, J. W. & Vitetta, E. S. *Fedn Proc.* **32**, 35-40 (1973).
12. Sherr, C. J. & Uhr, J. W. *J. exp. Med.* **133**, 901-920 (1971).
13. Sherr, C. J., Shenkin, I. & Uhr, J. W. *N.Y. Acad. Sci.* **190**, 250-367 (1971).
14. Putnam, F. W., Florent, G., Paul, C., Shinoda, T. & Shimizu, A. *Science* **182**, 287-291 (1973).
15. Ambler, R. P. *Meth. Enzym.* **25**, 262-272 (1967).

16. Winstead, J. A. & Wold, F. *J. biol. Chem.* **239**, 4212-4216 (1964).
17. Metzger, H. *Adv. Immun.* **12**, 57-116 (1970).
18. Spragg, B. P. & Clamp, J. R. *Biochem. J.* **114**, 57-64 (1969).
19. Shimizu, A., Putnam, F. W., Paul, C., Clamp, J. R. & Johnson, I. *Nature new Biol.* **231**, 73-76 (1971).
20. Tkacz, J. S. & Lampen, J. O. *Biochem. biophys. Res. Commun.* **65**, 248-257 (1975).
21. Waechter, C. J. & Lennarz, W. J. A. *Rev. Biochem.* **45**, 95-112 (1976).
22. Williamson, A. R., Singer, H. H., Singer, P. A. & Mosmann, T. R. *Biochem. Soc. Trans.* (in the press).
23. Schmeckpeper, B. J., Adams, J. M. & Harris, A. W. *FEBS Lett.* **53**, 95-98 (1975).
24. Jilka, R. L. & Pestka, S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5692-5696 (1977).
25. Singer, H. H., Gates, F. T. III, Kindt, T. J. & Williamson, A. R. *Eur. J. Immun.* (in the press).
26. Walsh, F. S. & Crumpton, J. J. *Nature* **269**, 307-311 (1977).
27. Parkhouse, R. M. E. & Cooper, M. D. *Immun. Rev.* **37**, 105-126 (1977).
28. Salsano, F., Froland, S., Natvig, J. B. & Michaelsen, T. E. *Scand. J. Immun.* **3**, 841-846 (1974).
29. Fu, S. M., Winchester, R. J. & Kundel, H. G. *J. Immun.* **114**, 250-252 (1975).
30. Goding, J. W. & Layton, J. E. *J. exp. Med.* **144**, 852-857 (1976).
31. Pernis, B., Brouet, J. C. & Seligmann, M. *Eur. J. Immun.* **4**, 776-778 (1974).
32. Tongegawa, S., Maxam, A. M., Tizard, R., Bernard, O. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1485-1489 (1978).
33. Schulman, M. J. & Köhler, G. *Nature* **274**, 917-919 (1978).
34. Marcu, K. B., Schibler, U. & Perry, R. P. *Science* **204**, 1087-1088 (1979).
35. Kessler, S. W. *J. Immun.* **115**, 1617-1624 (1975).
36. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
37. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46**, 83-88 (1974).
38. Kabat, E. A., Wu, T. T. & Bilofsky, H. *Variable Regions of Immunoglobulin Chains. Tabulation and Analyses of Amino Acid Sequences* (Bolt, Beranek and Newman, Massachusetts, 1976).
39. Sakano, H. *et al. Nature* **277**, 627-633 (1979).
40. Early, P. W., Davis, M. M., Kaback, D. B., Davidson, N. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 857-861 (1979).
41. Brack, C., Hiram, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* **15**, 1-14 (1978).
42. Honjo, T. *et al. Cell* **18**, 559-568 (1979).
43. Perry, R. P. & Kelley, D. E. *Cell* **18**, 1333-1339 (1979).
44. Early, P. *et al. Cell* (in the press).
45. Rogers, J. *et al. Cell* (in the press).

Fidelity of DNA replication catalysed *in vitro* on a natural DNA template by the T4 bacteriophage multi-enzyme complex

Urszula Hibner* & Bruce M. Alberts†

Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94143

*More than 50 copies of a Φ X174 DNA template can be made in 60 min in an *in vitro* DNA replication system consisting of seven purified replication proteins isolated from T4 bacteriophage-infected cells. By transfecting with the DNA products and assaying for the reversion of specific amber mutants, the high degree of base-pairing fidelity in this system is revealed; the *in vitro* system is also shown to respond to the mutagenic effect of Mn^{2+} and to display strong base-pair context effects on fidelity, as expected from *in vivo* studies.*

As a central genetic process, the synthesis of DNA proceeds with a very high fidelity in living cells. The average error frequency for base-pair substitutions has been estimated to be of the order of 10^{-8} and 10^{-10} errors per base pair replicated in T4 bacteriophage and *Escherichia coli* DNA replication, respectively¹. This fidelity is presumably due to Watson-Crick base-pairing recognitions; however, the fidelity expected from physicochemical knowledge of the stability of non-Watson-Crick base pairs and the occurrence of alternate tautomeric forms of the bases would predict 10^{-4} - 10^{-5} errors per base pair replicated².

DNA replication is catalysed by multi-enzyme complexes *in vivo*³⁻⁵. We have studied the fidelity of the replication reaction carried out *in vitro* by a multi-enzyme complex reconstituted from seven replication proteins coded for by T4 bacteriophage, each of which has been highly purified as described elsewhere⁶⁻⁸. This *in vitro* system contains the protein products of T4 genes 43 (DNA polymerase), 32 (helix destabilizing protein), 44 and 62 (a complex with DNA-dependent ATPase activity), 41 (a DNA-dependent GTPase and ATPase), 45 and 61. In general terms, the role of each of these proteins in the replication process has been defined; these results and many of the properties of the reconstituted DNA replication system have recently been reviewed^{8,9}. The complete system is capable of extensive DNA synthesis on a variety of double- and single-stranded DNA templates, closely mimicking the *in vivo* reaction with respect to the rate of chain elongation, RNA priming of Okazaki fragments and general structure of the replication fork^{8,13,14}. The studies presented here on the fidelity of DNA synthesis in this *in vitro* system, have been carried out using two different Φ X174

amber mutant DNAs as templates (*am3* and *am18*), and involve screening the product DNA molecules by transfection for the presence of wild-type Φ X174 revertants. The results demonstrate that remarkably few errors are made when the frequencies of specific base-pair substitutions are examined, allowing us to conclude that many of the elements of the DNA synthesis reaction that create the high fidelity observed *in vivo* also operate during the *in vitro* DNA synthesis catalysed by the reconstituted T4 system. Independent data supporting this conclusion have been obtained by Sinha and Haimes³⁵.

The fidelity assay

A highly sensitive measure of *in vitro* DNA replication fidelity is obtained by replicating the DNA of previously sequenced nonsense mutants of Φ X174 bacteriophage¹⁰ and assaying for revertants to wild type in the DNA synthesized^{8,11,12}. Beginning with either single-stranded or nicked double-stranded DNA circles as template, we have replicated Φ X amber mutant DNA by a 'rolling circle' mode, producing 20-100 copies in a brief incubation^{13,14}. As shown schematically in Fig. 1, the long double-stranded DNA product on the rolling circle tails is then cut to genome-sized fragments by treatment with a restriction endonuclease which recognizes only one site on the Φ X genome. The resulting linear molecules with 'sticky ends' are then circularized by T4 DNA ligase, and the infectivity of these DNA circles is measured by transfection into sup⁺ cells, in which the amber mutant grows [*E. coli* CQ2 (Su3⁺)]. Wild-type Φ X revertants are scored following transfection into sup⁻ cells (*E. coli* C), and the fraction of revertant DNA molecules calculated. The extensive synthesis obtained in our system yields mainly a double-stranded DNA product with both strands as newly synthesized

*Present address: Parasitologie Expérimentale, Institut Pasteur, 25 Rue du Dr Roux, 75015 Paris, France.

† To whom reprint requests should be addressed.

Table 1 Φ X174 DNA synthesized *in vitro* is fully infectious

DNA synthesized <i>in vitro</i>	Genome-sized DNA circles produced (% of total DNA)	Total DNA infectivity	Specific infectivity of DNA	Minimal estimate of specific infectivity of the newly synthesized DNA*
57.8 copies	29.5%	26.8%	0.91	0.86

The reaction mixture for DNA synthesis contained 10 mM Tris-acetate pH 7.8, 67 mM potassium acetate, 3.8 mM magnesium acetate, 0.5 mM dithiothreitol, 2 mM dCTP, 100 μ M dATP, 100 μ M dGTP, 20 μ M 3 H-dTTP (specific activity 200 c.p.m. μ mol $^{-1}$), 500 μ M rATP, 200 μ M each rUTP, rCTP and rGTP, and 0.5 μ g ml $^{-1}$ Φ X174 *am3* singly-nicked, double stranded circular DNA. The replication proteins present were 43 protein (3.3 μ g ml $^{-1}$), 32 protein (200 μ g ml $^{-1}$), 44/62 protein (18 μ g ml $^{-1}$), 45 protein (30 μ g ml $^{-1}$), 41 protein (20 μ g ml $^{-1}$) and 61 protein (0.16 μ g ml $^{-1}$). The initial reaction volume was 10 μ l. After a 45-min incubation at 37°C, the sample was diluted to 100 μ l with a reaction mixture containing all the components except the DNA. The reaction continued for 45 min at 37°C. The DNA synthesis was assayed by trichloroacetic acid precipitation of aliquots at 22, 45, 67 and 90 min; DNA was made at an accelerating rate until the last time point. The reaction was stopped by heating (68°C for 10 min). The DNA product was then digested with *Pst*I restriction endonuclease at 37°C (sevenfold over digestion). The reaction was stopped by heating at 68°C for 10 min. After cooling, rATP was added to 0.5 mM. The DNA was ligated with excess T4 DNA ligase overnight at 14°C. To favour formation of circles over dimerization, the concentration of DNA was kept low (2 μ g ml $^{-1}$). The DNA was extracted once with redistilled phenol equilibrated in 0.1 M Tris-HCl, pH 8.5 (This step is not necessary for subsequent transfection). For visualization in the electron microscope, the DNA was spread by a modification of Inman's procedure¹⁰, shadowed with platinum and analysed in a Philips 300 microscope. One-hundred randomly chosen molecules were traced. The transfection procedure used to measure infectivity was as described in Fig. 1 legend. Specific DNA infectivity = (% total infectivity)/(% DNA circles).

*Calculated by assuming that the input DNA template retained its full infectivity.

DNA¹⁴. This fact eliminates possible distortions in the assay due to cellular mechanisms which discriminate against newly synthesized DNA when repairing mismatched bases¹⁵.

Specific infectivity of DNA synthesized *in vitro*

Before studying the frequency of errors made in the *in vitro* DNA synthesis reaction on a Φ X amber mutant DNA template, we wished to establish that the product DNA is as active in infecting sup⁺ *E. coli* cells as the original template. Such a test is facilitated by the fact that extensive amounts of DNA can be synthesized by the T4 proteins *in vitro*. By using a dilution of an initial reaction mix as template for a second reaction, we have obtained an amount of DNA product corresponding to as much as 200 copies of the input circular template. In the experiment described in Table 1, we have synthesized 58 DNA copies on a Φ X double-stranded DNA template. The product DNA was processed as described above, and the proportion of genome-sized circular DNA in the preparation was estimated by electron microscopy. One hundred randomly chosen molecules were counted, 24 of which were unit-size circles. There was one double-genome length circle. The rest of the molecules were linear and of varying lengths. The genome-sized circles constituted 29.5% of the DNA mass, whereas 26.8% of the DNA was infectious (Table 1). We conclude that the specific infectivity of the DNA synthesized *in vitro* is indistinguishable from that of the original template. As the Φ X genome is 5,375 nucleotides long¹⁰, this result implies that less than about 2 errors are made per 10⁴ base-pair replications.

Fidelity of *in vitro* DNA synthesis at the Φ X174 bacteriophage *am3* site

We initially measured the frequency of *in vitro* base-pair substitutions at the Φ X174 *am3* mutant locus. As indicated in the Fig. 1 insert, the *am3* mutation changes a C·G base pair at position 587 to a T·A base pair¹⁶. This base-pair

change lies in the overlap of Φ X genes *E* and *D* (ref. 16), but it affects only the amino acid sequence in the gene *E* polypeptide (Fig. 1). In theory, any one of seven different single-base changes could change the polypeptide chain-terminating amber codon in gene *E* into a codon which specifies some amino acid; however, some of these would change the amino acid specified in the gene *E* and/or gene *D* proteins in such a way as still to produce defective phage. The number of changes in the *am3* codon that could produce viable phage is not known. However, as our *in vitro* reaction conditions strongly favour reversion to the true wild type (see below), we believe that we are mainly measuring the frequency of T·A to C·G transitions at position 587 in our assays. This assumption is supported by the DNA sequencing of similarly produced *in vitro* revertant DNA molecules which has been carried out in other laboratories (ref. 35 and L. A. Loeb, personal communication).

Due to the relatively high *in vivo* reversion frequency of Φ X mutants, the population of *am3* DNA used as template in our experiments is contaminated with what seems to be wild-type revertants, at a level of about 3×10^{-6} . To increase the number of *in vitro* revertants above background, we have strongly biased the input concentrations of deoxyribonucleoside triphosphates (dNTPs) to favour the T·A to C·G change necessary to produce wild-type revertants. As dGTP competes with dATP to make a G·T mispair¹², a high concentration of dGTP relative to dATP will increase the probability of the G·T mispairs needed to obtain revertants to wild type when synthesizing the viral DNA strand (Fig. 1); similarly, a high concentration of dCTP relative to dTTP will increase the probability of the C·A mispairs needed to obtain revertants to wild type when synthesizing the complementary

Table 2 The frequency of an AT to GC transition depends on the mispaired intermediate

dNTP bias during the <i>in vitro</i> reaction	$\frac{dCTP}{dTTP} = \frac{100}{1}$	$\frac{dGTP}{dATP} = \frac{51}{1}$
(1) No. of infectious Φ X molecules tested	5.9×10^9	1.6×10^9
(2) No. of infective centres expected if all DNA was wild-type	5.3×10^5	1.25×10^6
(3) No. of infective centres found	7	74
(4) No. of infective centres due to wild-type contaminants in template	2	5
(5) Level of <i>in vitro</i> -produced wild-type revertants in <i>am3</i> populations	9.4×10^{-6}	5.5×10^{-5}
(6) Calculated error frequency	1.9×10^{-7}	2.2×10^{-6}
	(C pairing with A)	(G pairing with T)

The reaction conditions were as described in Table 1 legend, except that the reaction volumes were 100 μ l (dCTP/dTTP bias) or 50 μ l (dGTP/dATP bias) and nicked double-stranded Φ X *am3* replicative form DNA was used as template at 4 μ g ml $^{-1}$ (wild-type revertant level of 4×10^{-6}). In the experiment with a dGTP/dATP bias, the concentrations of dNTPs were 1 mM dGTP, 20 μ M dATP, 100 μ M dCTP and 100 μ M 3 H-dTTP (specific activity of 200 c.p.m. μ mol $^{-1}$). The transfections were carried out with both *E. coli* CQ2 (su3) and *E. coli* C (su⁻), as described in Fig. 1. To control for possible inhibition of expression of the small amount of wild-type revertants present by the large excess of Φ X *am3* DNA, the su⁻ cells were also transfected with a mixture of the *in vitro* synthesized DNA with known amounts of wild-type DNA. Our results are thereby corrected for possible alterations of transfection efficiencies due to any component present in the *in vitro* DNA product. The use of dCTP/dTTP or dGTP/dATP bias in the reaction is expected to favour the AT→GC transition at position 587 (leading to wild-type revertants), as well as TA→CG transition at position 586. The latter change would result in changes in both the gene *E* (Trp→Gln) and the gene *D* (Val→Ala) proteins. It is not known whether the mutant Φ X produced from *am3* by the T·A→C·G transition at position 586 is viable. This point is, however, not crucial to our argument, because at worst we may be somewhat underestimating the replication fidelity in our *in vitro* system (that is, the reversion rate observed could possibly be a sum of reversions occurring at two neighbouring loci: positions 586 and 587). These data differ somewhat from those reported earlier⁸, in that we previously underestimated the error frequencies for a G/A bias due to a technical problem.

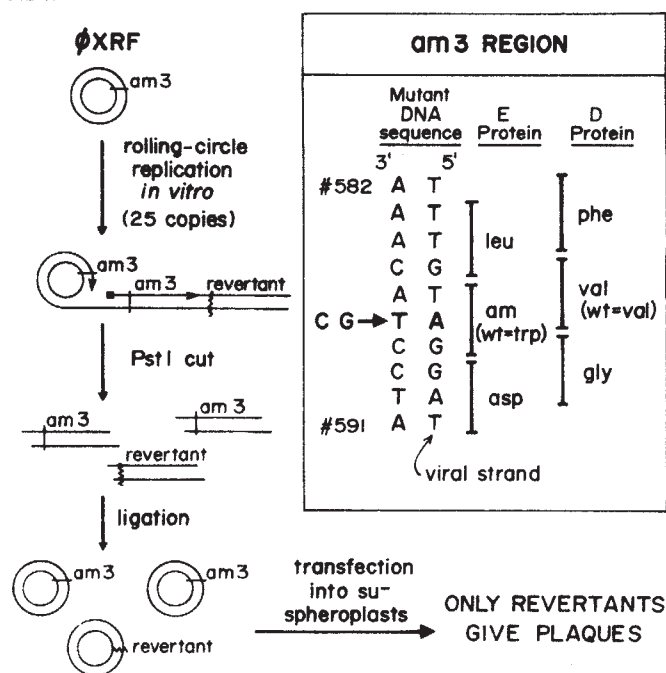


Fig. 1 Schematic summary of an infectivity assay for replication fidelity. The two arrows on the rolling circle denote the position and direction of the DNA polymerase molecule on the leading strand (intact circular template) and lagging strand (linear template displaced from circle) of the replication fork, respectively³. A detailed physical characterization of the products made has been reported elsewhere^{13,14}. The DNA from an amber mutant of bacteriophage ΦX174 is replicated *in vitro* in a buffer containing 10 mM Tris-acetate, pH 7.8, 67 mM potassium acetate, 3.8 mM magnesium acetate (unless otherwise stated), 0.5 mM dithiothreitol, 0.5 mM rATP and 0.2 mM each of rGTP, rUTP and rCTP. The concentrations of dNTPs and ΦX DNA (double- or single-stranded) present varied, as described in each case. The concentrations of T4 replication proteins used are given in Table 1 legend. The incubations were at 37°C for 30–90 min. The time course and the extent of DNA synthesis were assayed by trichloroacetic acid precipitation of aliquots. The reactions were stopped either by heating the samples at 68°C for 10 min or by phenol extraction. In all cases, contamination of dCTP by dUTP in the commercial dCTP preparations was eliminated by the use of purified dCTP or by treatment of reaction mixtures with dUTPase. The transformation procedure used is a modification of a method of Kushner³⁴. A volume of 100 ml of *E. coli* CQ2 (Su3⁺) is grown in L broth to a density of $2-3 \times 10^8$ cells ml⁻¹. The cells are spun and washed once at 4°C in 50 ml of 10 mM MOPS pH 7.0, 10 mM RbCl. The washed cell pellet is resuspended in 50 ml of 0.1 M MOPS, pH 6.5, and 10 mM RbCl. A 0.5-ml volume of 5 M CaCl₂ is added and the cells are incubated at 0°C for 30 min. The cells are pelleted and resuspended gently at 4°C in 10 ml of 0.1 M MOPS pH 6.5, 10 mM RbCl and 50 mM CaCl₂. Aliquots of 0.2 ml are distributed to chilled glass tubes to which 3 μl dimethyl sulphoxide is added. The DNA is added to each tube in a volume of 10 μl or less. After 30 min at 0°C the samples are heated to 45°C for 30 s. L broth (1 ml containing 1 mM CaCl₂) is then added and the samples are incubated at 37°C for 10 min. Top agar containing indicator bacteria (either *E. coli* CQ2 or *E. coli* C) is added and the mixture poured onto bottom agar plates. The ΦX174 plaques are scored after about 3 h at 37°C. In controls with wild-type ΦX phage DNA, the transfection efficiencies for double-stranded ΦX circles were between 2×10^{-4} and 10^{-3} infective centres per DNA molecule for *E. coli* C and between 2×10^{-5} and 5×10^{-5} infective centres per DNA molecule for *E. coli* CQ2 cells; the true fraction of wild-type ΦX revertants is calculated by reference to calibrations with traces of wild-type DNA added as internal standards (see Table 2 legend). Single-stranded circular ΦX DNA molecules and linear double-stranded molecules do not contribute to the assay, as each is at least two orders of magnitude less infectious than double-stranded circles (data not shown).

DNA strand. Thus, in theory at least, the triphosphate bias not only makes these replication errors detectable, but also allows one to determine the types of mispair involved in creating the error.

Typical results of an experiment carried out with a high dNTP bias of each type are shown in Table 2. Comparison of the number of infective centres found (row 3) with the number expected from the level of pre-existing revertants in the template (row 4) indicates that a clearly detectable revertant

population is being produced during these *in vitro* DNA synthesis reactions. The fact that the detection of such *in vitro* revertants requires the expected triphosphate bias suggests that we are indeed observing reversion to true wild type. In row 6 of Table 2, we have assumed a linear dependence of base-pairing errors on the bias (see below) to calculate the error frequency expected for an equimolar dNTP concentration of 100 μM each. As indicated, it seems that a C pairs with A much less often than a G pairs with T at position 587. The fidelity of the multi-enzyme T4 *in vitro* DNA synthesis system is impressive; the error rates of 2×10^{-7} to 2×10^{-6} are 100–1,000-fold lower than those estimated for DNA polymerases alone on a ΦX DNA template (ref. 17).

Both single-stranded and double-stranded ΦX DNA circles can be used as a template for rolling circle DNA replication *in vitro*^{13,14}. In both cases, the reaction is a highly processive one, with very few secondary replication forks established on the rolling circle tails¹⁴. Therefore, with a single-stranded viral DNA circle as template, all but the very first copy of the complementary DNA strand will be synthesized at the leading side of a standard replication fork, and all of the viral strand will be synthesized at the lagging side of this fork (Fig. 1; see refs 13, 14 for details). In contrast, because we rely on nonspecific nicking of a double-stranded template DNA for initiation of rolling-circle replication, there is an equal probability that either strand (viral or complementary) will be made on each side of the fork on a double-helical template. From the fact that the estimated error frequencies at am3 for both viral (dGTP/dATP bias) and complementary (dCTP/dTTP bias) strand syntheses are the same irrespective of the type of template used (data not shown), we conclude that the fidelities of leading and lagging strand synthesis at the fork must be very similar, if not identical.

The fidelity of *in vitro* DNA synthesis is reduced in the presence of Mn²⁺

Substitution of manganese for magnesium is known to be mutagenic both *in vivo*^{18,19} and *in vitro* (for review see ref. 20). *In vivo*, the mutagenic effect of manganese on T4 DNA

Table 3 Substitution of manganese for magnesium ions during the *in vitro* DNA synthesis lowers the replication fidelity

dNTP bias during the <i>in vitro</i> reaction	dCTP dTTP = 20	dGTP dATP = 20
(1) No. of infectious ΦX molecules tested	3.9×10^8	7.5×10^8
(2) No. of infective centres expected if all DNA was wild type	1.8×10^5	1.4×10^5
(3) No. of infective centres found	8	98
(4) No. of infective centres due to wild-type contaminants in template	0.7	0.6
(5) Level of <i>in vitro</i> -produced wild-type revertants in am3 populations	4.1×10^{-5}	7.0×10^{-4}
(6) Calculated error frequency	4.1×10^{-6}	7.0×10^{-5}

The reaction conditions were the same as for the experiment in Table 2, except that ΦX174 single-stranded DNA ($5 \mu\text{g ml}^{-1}$) was used as template (background wild-type revertants of 4×10^{-6}). The 50-μl reactions contained either 800 μM dCTP, 40 μM ³H-dTTP (specific activity 200 c.p.m. pmol⁻¹) and 100 μM each of dATP and dGTP (C/T bias), or 400 μM dGTP, 20 μM dATP and 100 μM each of dCTP and ³H-dTTP (G/A bias). The magnesium salt in our normal reaction mixture was substituted with MnCl₂ at 2.2 mM or 1.2 mM for dCTP/dTTP and dGTP/dATP biased reactions, respectively. In other experiments, the exact level of Mn²⁺ present was shown not to affect the fidelity and the different levels used here reflect an attempt always to have about 1 mM non-triphosphate-bound Mn²⁺ present. The reactions were stopped by extractions with re-distilled phenol followed by ethanol precipitation. The restriction enzyme digestion and ligation were carried out in Mg²⁺-containing buffers.

Table 4 Site specificity of mutation rates

Type of <i>in vitro</i> reaction		Fidelity (fraction of errors)	
dNTP bias	Divalent cation	Mutant <i>am3</i>	Mutant <i>am18</i>
		³ AAACA CCTA ⁵ , ⁵ TTTGT AGGAT ³	³ CTGCA TCTTC ⁵ , ⁵ GACG TAAAG ³
C>T	Mg ²⁺	$3.3 \pm 0.3 \times 10^{-7}$	$2.0 \pm 0.5 \times 10^{-6}$
G>A	Mg ²⁺	$2.2 \pm 0.3 \times 10^{-6}$	$9.1 \pm 2.9 \times 10^{-6}$
C>T	Mn ²⁺	$4.6 \pm 1.1 \times 10^{-6}$	$5.8 \pm 2.3 \times 10^{-6}$
G>A	Mn ²⁺	$5.0 \pm 0.9 \times 10^{-5}$	$1.3 \pm 0.2 \times 10^{-5}$

The reactions were carried out as described in the legends to Tables 1–3. The results of 2–10 experiments of each type are presented as the mean value plus or minus the standard error. Double-stranded Φ X174 replicative form DNA which had been singly nicked with pancreatic DNase I, or single-stranded Φ X174 viral DNA was used as the template in duplicate reactions. The restriction enzyme used for digestion of the rolling circle tails was *Pst*I, *Ava*I or *Ava*II. Either T4 DNA ligase or *E. coli* ligase was used in the presence of 0.5 mM rATP or 26 μ M NAD, respectively, to create the double-stranded DNA circles used for transfections (see Fig. 1). The Φ X174 amber DNA preparations used as template contained 4×10^{-6} wild-type revertants for *am3*, and 2×10^{-6} wild-type revertants for *am18*. The sequence of each amber mutant DNA is shown, with the base pair in the box requiring a T·A→C·G transition to produce a wild-type revertant. For both mutants, the viral strand sequence is listed below the complementary strand sequence. The test for revertants at the *am18* site was carried out at 42°C, to prevent growth of a pseudo-revertant produced by a G→C transversion at position 25 (ref. 25).

synthesis is expressed mainly as an increase in the frequency of transitions¹⁸, both G·C→A·T and A·T→G·C transitions being enhanced by the order of 10–100-fold.

Our *in vitro* T4 DNA replication system is capable of using manganese as a divalent cation, although the range of concentrations permitting extensive synthesis is narrow. At optimal Mn²⁺ concentrations (1–2 mM), the rate of DNA synthesis is two or three times slower than in Mg²⁺ (data not shown). As shown in Table 3, the fidelity of DNA synthesis catalysed by the T4 replication proteins is lower in manganese- than in magnesium-containing buffers, as revealed by the enhanced frequency of revertants observed at the *am3* locus. The error frequencies on both the viral strand (dGTP/dATP bias) and the complementary strand (dCTP/dTTP bias) are enhanced about 20-fold in these mutagenic conditions.

The production of revertants *in vitro* increases linearly with dNTP bias

Although results obtained in the presence of Mn²⁺ ions are less relevant to the process of DNA synthesis *in vivo* than results obtained in Mg²⁺, the higher frequency of revertants observed in Mn²⁺ allowed us to test directly a crucial assumption made in the fidelity calculations presented previously in Tables 2 and 3, namely that the probability of misincorporation of an incorrect dNTP is directly proportional to the ratio of the concentration of incorrect to correct substrate. In Fig. 2, we plot the compiled results of several experiments carried out in Mn²⁺ which were designed to test the effect of triphosphate bias on the frequency of *am3* revertant production. The data points fit satisfactorily to the lines drawn with a slope of 1.0 on this log–log plot, a statistical analysis indicating actual best fit slopes of 0.91 (± 0.18) and 0.89 (± 0.12) for the dCTP/dTTP and the dGTP/dATP bias experiments, respectively. We conclude that both types of bias cause a simple linear increase in the appearance of revertants, as assumed.

Complications due to exonuclease ‘proofreading’

Interpretation of the linear relationship between errors and triphosphate bias in Fig. 2 is complicated by the fact that misincorporated nucleotides tend to be preferentially removed by the 3'→5' proofreading exonuclease activity of DNA polymerase²¹. Because of the constant competition between polymerization and exonuclease removal, the fidelity of DNA

synthesis at each template position is strongly influenced by the rate of addition of the next dNTP normally incorporated in the DNA chain, and the error is only fixed after addition of this subsequent nucleotide²². For example, when synthesizing the wild-type viral strand DNA sequence on the *am3* template, any G incorporated at 587 will be ‘paired’ with a T on the template, and therefore be very liable to exonuclease removal. Fersht has recently shown¹¹ that, in a Φ X RF→SS DNA reaction, *E. coli* DNA polymerase III holoenzyme creates viral strand revertants at this sequence in proportion to [dGTP/dATP]². This is the expected result if the rate of addition of the next nucleotide, a G at position 588, increases linearly with dGTP concentration in the range tested (0.008–1.0 mM), thereby reducing the amount of time available for the exonuclease function of the polymerase to remove a misincorporated G residue at position 587 (ref. 12). However, this effect should disappear when dGTP concentrations are sufficient to give close to maximal polymerization rates.

Extrapolating to our system, the results of Goodman and co-workers with T4 DNA polymerase²² would predict an effect of dGTP concentration on exonuclease proofreading at position 587 only when this concentration drops below ~100 μ M. In fact, we obtained a [dGTP]/[dATP] dependence of error frequencies in Fig. 2 only because all our dNTP concentrations are usually kept at a high level (100 μ M each); thus, all the dGTP concentrations used to create the dGTP/dATP biases in Fig. 2 are sufficient to maximize the rate of G addition at position 588.

As predicted by the data of Fersht¹² and Clayton *et al.*²², there is a marked increase in the fidelity of synthesizing the *am3* viral strand in our system when the G to A ratio is kept at 1.0 and the dNTP concentrations are reduced from 100 μ M to 20 μ M. In fact, to measure the reversion frequencies on the complementary strand (T→C transitions) at low dCTP/dTTP bias in Fig. 2, it was necessary to lower the nonbiased dGTP and dATP concentrations. At higher levels of dGTP, the preferential A→G reversions on the viral strand prevent detection of the T→C transitions being tested (see Fig. 2 legend).

The fidelity of DNA synthesis *in vitro* is site dependent

The frequency of a given base-pair mutation *in vivo* is strongly dependent on the location of that base pair in the context of neighbouring DNA sequences; for the same base-pair change, mutation frequencies can differ by several orders of magnitude^{23,24}.

We have compared the reversion frequencies of two different Φ X174 mutants, *am3* (gene *E*) and *am18* (gene *A*), after *in vitro* DNA replication by the T4 multi-enzyme complex. The *am18* mutant is at position 23 of the Φ X DNA sequence²⁵, in the region of overlap of genes *A* and *B*. The DNA sequences around nucleotides 23 (*am18*) and 587 (*am3*) are compared at the top of Table 4. As indicated (boxed base pairs), reversion to wild type requires a T·A to C·G transition for both mutants.

As reported in Table 4, the *in vitro* reversion frequencies for *am3* and *am18* are quite distinct. The frequency of T→C transitions is about sevenfold higher for *am18* than for *am3*, whereas the A→G transition frequency is higher by about a factor of 4 at the *am18* site. At both sites, we obtain a higher number of Φ X revertants when using a dGTP/dATP bias (favouring G·T mispairs) than when using a dCTP/dTTP bias (favouring C·A mispairs), suggesting that the G·T mispair may more easily escape discrimination *in vitro*. This result may be related to the finding that the G·T base pair is much more stable than the A·C base pair, when tested in synthetic polynucleotide helices (J. Fresco, personal communication).

In Table 4, the substitution of manganese for magnesium during the *in vitro* DNA synthesis leads to an increase in error rate in all cases. The extent of the increase caused by Mn²⁺,

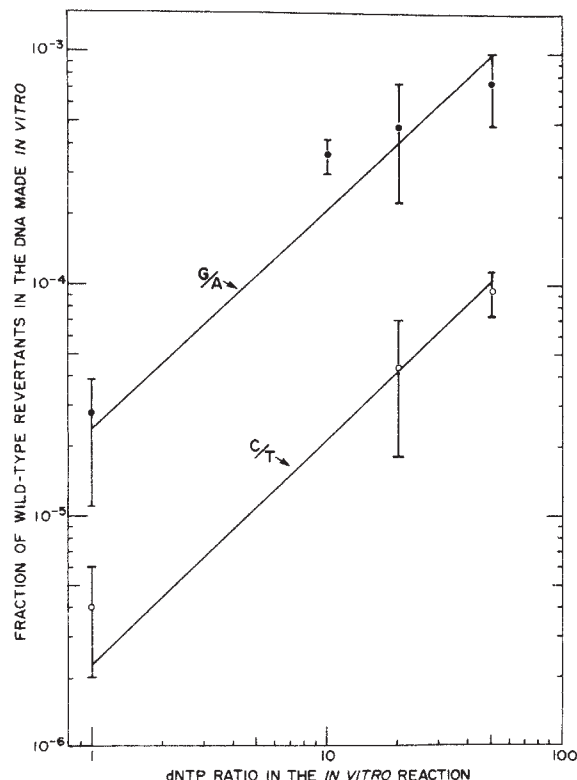


Fig. 2 Linear dependence of *in vitro* revertant frequencies on dNTP bias. The lines shown are the best fit to a slope of 1.0 on this log-log plot; an appropriately weighted statistical analysis gives actual slopes of 0.91 (± 0.18) and 0.89 (± 0.12) for the dCTP/dTTP and the dGTP/dATP bias experiments, respectively. The reactions were carried out on a $\Phi\chi$ *am3* single-stranded DNA template as described in Tables 1 and 3 legends. For the dGTP/dATP bias experiments, the dNTP concentrations were: 100 μ M dCTP, 100 μ M dTTP, 100 μ M dGTP, 100 μ M dATP ($G/A=1$); 100 μ M dCTP, 100 μ M dTTP, 200 μ M dGTP, 20 μ M dATP ($G/A=10$); 100 μ M dCTP, 100 μ M dTTP, 400 μ M dGTP, 20 μ M dATP ($G/A=20$); 100 μ M dCTP, 100 μ M dTTP, 1 mM dGTP, 20 μ M dATP ($G/A=50$). For the dCTP/dTTP bias, the dNTP concentrations were: 100 μ M dCTP, 100 μ M dTTP, 20 μ M dGTP, 20 μ M dATP ($C/T=1$); 400 μ M dCTP, 100 μ M dTTP, 100 μ M dGTP, 100 μ M dATP ($C/T=20$); 1 mM dCTP, 20 μ M dTTP, 100 μ M dGTP, 100 μ M dATP ($C/T=50$). Reactions were stopped while the amount of DNA product was still increasing linearly, and we estimate that not more than half of the dNTP present in lowest concentration had been used up. Note that by changing the bias of one pair of triphosphates, we also affect the ratios of these triphosphates to the unbiased pair. However, because of the way the experiments were carried out, these other ratios did not change by the same factor. For example, changing the dGTP/dATP ratio from 1- to 50-fold changed the dGTP/dTTP and dGTP/dCTP ratios only 10-fold. These two particular changes could possibly lead to the production of pseudo-wild-type revertants at the *am3* locus by enhancing transversions; however, the linear correlation between dGTP/dATP ratio and the production of revertants *in vitro* suggests that most of the viable revertants are true wild type.

however, is very different for the different sites. The greatest effect was observed for the apparent A $\rightarrow$ G substitution at the *am3* site (about 20-fold higher errors in Mn^{2+} than in Mg^{2+}), whereas a 1.5-fold increase was found in the A $\rightarrow$ G substitutions at the *am18* site.

We find that adding increasing amounts of magnesium to manganese-containing reaction mixtures (up to a 10:1 Mg^{2+} to Mn^{2+} ratio) does not reduce the high error frequency characteristic of the manganese reaction (data not shown). As the binding constants of Mn^{2+} and Mg^{2+} for dNTPs are very similar²⁶, the absence of a strong Mg^{2+} competition effect in these fidelity experiments seems incompatible with the idea that manganese effects are exerted only through its complex with the dNTPs. Rather, the mutagenic activity of manganese is likely to be due largely to its interaction with the DNA polymerase or with other replication proteins.

Is *in vitro* DNA synthesis as faithful as the *in vivo* process?

The average error rate for point mutations measured for T4 bacteriophage *in vivo* is about 2×10^{-8} per base-pair replication¹. However, when the reverse mutation rates for a number of rII loci are compared, it becomes clear that reversions occurring by exactly the same pathway (for example, a G $\cdot$ C $\rightarrow$ A $\cdot$ T transition) vary in frequency by several orders of magnitude (for example mutant *uvl* gives 5×10^{-10} , mutant *uv19* gives 8×10^{-8} and mutant *uv102* gives 10^{-6} revertants per replication)²³. Our estimates of overall *in vitro* frequencies for a T $\cdot$ A to C $\cdot$ G transition are about 10^{-6} and 10^{-5} revertants per base-pair replication for the $\Phi\chi$ 174 *am3* and *am18* loci, respectively. Thus, although impressively accurate, the *in vitro* replication fidelity seems to be inferior to the *in vivo* precision observed in T4 bacteriophage-infected cells.

Several explanations for the above discrepancy should be considered. (1) It is possible, although unlikely, that the two $\Phi\chi$ 174 mutations whose *in vitro* reversion has been studied by us represent unusually mutable sites for the T4 replication apparatus, and thus that the data in Table 4 cannot be assumed even to approximate the average of the spectrum of mutation frequencies. There is some suggestion to this effect in ref. 35. (2) The *in vivo* T4 mutation rate estimates may include contributions to fidelity from post-replication repair mechanisms which do not function in our *in vitro* system. In particular, a repair mechanism linked to DNA methylation is in principle capable of recognizing errors on newly synthesized DNA, and there is evidence that such a repair pathway increases the fidelity of *E. coli* DNA replication by a factor of 100–1,000 (ref. 15). The presence of a similar repair pathway operating in T4-infected cells could be tested for by measuring mutation frequencies for a T4 *dam*⁻ phage mutant²⁷ grown on an *E. coli* *dam*⁻ host¹⁵. (3) The conditions in which our *in vitro* DNA synthesis is carried out have not been specifically optimized to obtain maximum fidelity. Although fidelity could be increased by lowering dNTP concentrations^{12,22}, this would reduce the rate of *in vitro* replication fork movement far below *in vivo* levels⁸. However, it is possible that fidelity could be increased without slowing polymerization rates by altering ionic conditions, adding polyamines or increasing ATP to ADP ratios²⁸. (4) The present T4 *in vitro* DNA replication system may be lacking as yet unidentified T4 or host gene products which are essential only for maintenance of maximum replication fidelity. There is good genetic evidence that such special proteins exist for the DNA replication system of uninfected *E. coli*²⁹, as well as indirect evidence that additional T4 proteins directly feed dNTPs into the *in vivo* T4 DNA replication complex^{30,31}.

Conclusions

Despite the above uncertainties, the major point of our current findings is that the rapid and extensive DNA replication observed in the present T4 *in vitro* system is highly faithful. It also responds to the mutagenic effect of Mn^{2+} , and displays the strong base-pair context effects on fidelity expected from *in vivo* studies. Genetic analyses reveal that an alteration in any one of the seven T4 DNA replication proteins used in our *in vitro* system can increase mutation rates^{32,33}. Future *in vitro* studies using this system should thus allow elucidation of the detailed mechanisms of many of the fidelity-generating steps in DNA replication.

This work was supported by USPHS grant GM 24020 from the National Institute of General Medical Sciences and a postdoctoral fellowship from the Damon Runyon-Walter Winchell Cancer Fund to U.H. We thank Mei Lei Wong for help with electron microscopy, Y. Shlomai and A. Kornberg for dUTPase, P. Weisbeek for $\Phi\chi$ 174 *am18* bacteriophage, and our colleagues at UCSF for advice on the manuscript.

Received 22 October 1979; accepted 26 March 1980.

1. Drake, J. W. *Nature* **221**, 1132 (1969).
2. Topal, M. D. & Fresco, J. R. *Nature* **263**, 285–289 (1967).
3. Alberts, B. & Sternglanz, R. *Nature* **269**, 655–661 (1977).
4. Wickner, S. H. A. *Rev. Biochem.* **47**, 1163–1191 (1978).
5. Kornberg, A. *Cold Spring Harb. Symp. quant. Biol.* **43**, 1–9 (1979).
6. Morris, C. F., Hama-Inaba, H., Mace, D., Sinha, N. K. & Alberts, B. J. *biol. Chem.* **254**, 6787–6796 (1979).
7. Morris, C. F., Moran, L. A. & Alberts, B. M. J. *biol. Chem.* **254**, 6797–6802 (1979).
8. Liu, C.-C., Burke, R. L., Hibner, U., Barry, J. & Alberts, B. M. *Cold Spring Harb. Symp. quant. Biol.* **43**, 469–487 (1979).
9. Alberts, B. et al. in *Proc. P & S biomed. Sci. Symp.* (ed. Vogel H.J.) 31–63 (Academic, New York, 1977).
10. Sanger, F. et al. *Nature* **265**, 687–695 (1977).
11. Weymouth, L. A. & Loeb, L. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1924–1928 (1978).
12. Fersht, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4946–4950 (1979).
13. Morris, C. F., Sinha, N. K. & Alberts, B. M. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4800–4804 (1975).
14. Sinha, N. K., Morris, C. F. & Alberts, B. M. J. *biol. Chem.* **255**, 4290–4303 (1980).
15. Radman, M. et al. *Cold Spring Harb. Symp. quant. Biol.* **43**, 937–946 (1979).
16. Barrell, B. G., Air, G. M. & Hutchison, C. A. III *Nature* **264**, 34–41 (1976).
17. Loeb, L. A. et al. *Cold Spring Harb. Symp. quant. Biol.* **43**, 921–928 (1979).
18. Orgel, A. & Orgel, L. E. *J. molec. Biol.* **14**, 453–457 (1965).
19. Ripley, L. S. *Molec. gen. genet.* **141**, 23–40 (1975).
20. Mildvan, A. S. & Loeb, L. A. *CRC crit. Rev. Biochem.* **6**, 219–244 (1979).
21. Brutlag, D. & Kornberg, A. J. *biol. Chem.* **247**, 241–248 (1972).
22. Clayton, L. K., Goodman, M. F., Branscomb, E. W. & Galas, D. J. *J. biol. Chem.* **254**, 1902–1912 (1979).
23. Drake, J. W. *The Molecular Basis of Mutation* (Holden Day, San Francisco, 1970).
24. Coulondre, C. & Miller, J. M. J. *molec. Biol.* **117**, 525–575 (1977).
25. Smith, M. et al. *Nature* **265**, 702–705 (1977).
26. Mohan, M. S. & Rechnitz, G. A. *Archs Biochem. Biophys.* **162**, 194–199 (1974).
27. Hattman, S., van Ormondt, H. & de Waard, A. J. *molec. Biol.* **119**, 361–376 (1978).
28. Kurland, C. G. *Biophys. J.* **22**, 373–392 (1978).
29. Cox, E. C. A. *Rev. Genet.* **10**, 135–156 (1976).
30. Wovcha, M. G., Tomich, P. K., Chiu, C.-S. & Greenberg, G. R. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2196–220 (1973).
31. Reddy, G. P. V., Singh, A., Stafford, M. E. & Mathews, C. K. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3152–3156 (1977).
32. Mufti, S. *Virology* **94**, 1 (1979).
33. Mufti, S. & Bernstein, H. J. *J. Virol.* **14**, 860–871 (1974).
34. Kushner, S. R. in *Proc. int. Symp. Genetic Engineering* (eds Boyer, H. W. & Nicosia, J.) 17–23 (Elsevier, Amsterdam, 1978).
35. Sinha, N. K. & Haimes, M. D. in *Mechanistic Studies of DNA Replication and Genetic Recombination* (CN-UCLA Symp., in the press).

LETTERS

The Centaurus I cluster of galaxies —An extreme case of contamination?

J. R. Lucey*, R. J. Dickens† & J. A. Dawe‡

*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK

†Anglo-Australian Observatory, Private Bag, Coonabarabran, New South Wales 2857, Australia

‡UK Schmidt Telescope Unit (Royal Observatory Edinburgh), Private Bag, Coonabarabran, New South Wales 2857, Australia

Although the virial theorem is often used to calculate a system's mass from its size and velocity dispersion, there is considerable uncertainty over the result because of the problem of defining the true membership of the system. Recent analysis¹ of hypothetical galaxy systems in model universes has indicated that this local contamination can cause large errors in virial mass estimates for rich clusters. Non-member galaxies projected in the line of sight onto clusters are not easily recognized, unless they have redshifts significantly different from the cluster mean redshift. Contaminating galaxies more local to the cluster (for example, those with a small redshift difference) will be hidden in the large intrinsic motions of the cluster members. Other difficulties associated with application of the virial theorem include possible substructure within the clusters and non-virial motions (for example, infall or expansion). Such effects obscure the physical meaning of a derived virial mass. Here we emphasize the reality of these problems, as evidenced by the resolution of the Centaurus I cluster of galaxies into two distinct velocity systems, based on extensive new redshift data².

The Centaurus I cluster of galaxies has been studied frequently in recent years. Sandage³ found, from 14 redshifts, ($n=14$), a heliocentric mean cluster velocity ($\bar{V}$) of $3,360 \text{ km s}^{-1}$ and a line-of-sight velocity dispersion (σ_v) of 673 km s^{-1} . Dawe et al.⁴, from $n=61$, found $\bar{V}=3,510 \text{ km s}^{-1}$ with $\sigma_v=870 \text{ km s}^{-1}$. Vidal and Wickramasinghe⁵, from $n=25$, found $\bar{V}=3,507 \text{ km s}^{-1}$ with $\sigma_v=945 \text{ km s}^{-1}$. The brightest galaxy in the cluster, NGC4696, is the radio source⁶ PKS1245–41. This is also the centre of an extended X-ray source, in which the 7-keV Fe XXV, XXIV emission feature has been observed^{7,8}.

We have obtained new radial velocities with the image photon counting system on the Anglo-Australian telescope. In addition, the spectral scans obtained by Dawe et al.⁴ have been remeasured with improved precision. These data have been combined with published results to yield redshifts for 164 galaxies within an UK Schmidt Telescope field centred on the cluster. The distribution of the velocities for 149 galaxies with

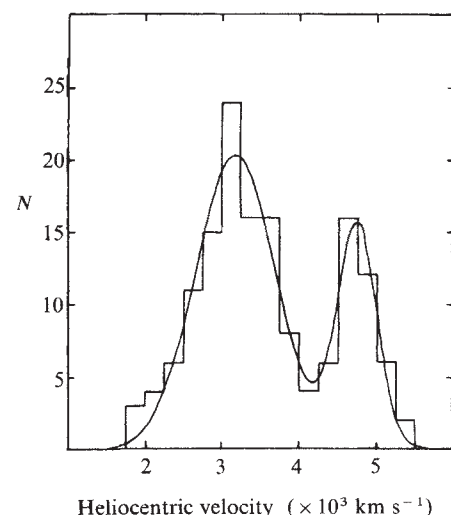


Fig. 1 Velocity distribution for galaxies on UKSTU plate, B1340 centred on NGC4696.

$V < 6,000 \text{ km s}^{-1}$ is shown in Fig. 1. The histogram is clearly bimodal, with the very low probability ($P < 0.001$) of such a form originating from a single normal distribution. The 'best-fit' double normal distribution, also shown in Fig. 1, has the parameters $n_1=106$, $\bar{V}_1=3,179 \text{ km s}^{-1}$, $\sigma_v=517 \text{ km s}^{-1}$; $n_2=40$, $\bar{V}_2=4,738 \text{ km s}^{-1}$, $\sigma_v=255 \text{ km s}^{-1}$. (The dispersions are reduced to 507 and 235 km s^{-1} respectively, after correction for the measuring error of 100 km s^{-1} .) The spatial distributions of the two velocity systems are shown in Fig. 2.

If both clusters were at their respective Hubble distances, the expected difference in their distance moduli ($\Delta\mu$) would be 0.9 mag. If this were the case, the observed number of galaxies to the same absolute magnitude limit would indicate that the clusters have equal populations (richness), a surprising result in view of the large difference in their velocity dispersions. Direct comparison of the cluster luminosity functions indicates $\Delta\mu=1.0 \pm 0.1 \text{ mag}$, which is in good agreement with the hypothesis. Alternatively if the clusters were physically related, then their expected similar distance implies the lower velocity system to be about 2.5 times as rich as the other system. Allowing for the difference in richness, comparison of the luminosity functions indicates $\Delta\mu=0.0 \pm 0.2 \text{ mag}$. Although the fit is poorer than the previous case, the magnitude shift is also consistent with the hypothesis. A definitive test is not possible at present because the form and variety of cluster luminosity functions makes richness and distance effects difficult to se-

Table 1 Galaxy groups near Centaurus I

Group name	<i>n</i>	$v_p(\text{km s}^{-1})$	$\sigma(\text{km s}^{-1})$	$r(\text{NGC4696})$
NGC4373	10	3,320	199	4.7°
NGC5011	8	3,251	160	4.7°
NGC5090	5	3,828	412	6.4°
ESO323-G77*	5	4,817	156	3.4°

*See ref. 13.

parate. Both interpretations are quite plausible, the former because superposition of systems on galaxy clusters is not unusual⁹⁻¹¹ and the latter because the observed velocity difference, being less than the r.m.s. escape velocity could arise from infall of the two clusters within a bound system. The growth of clusters by the amalgamation of subunits is a characteristic phenomenon in *N*-body studies¹³. In either case the existence of two systems will give a substantially reduced total virial mass.

There is also evidence that other velocity systems may be present. Around the main concentrations, four distinct galaxy groups are apparent, details of which are given in Table 1. The velocities of the NGC4373 and NGC5011 groups suggest that they are associated with the lower velocity clusters whereas the ESO 323-G77 group has a mean velocity close to that of the higher velocity cluster. Any similar groups projected on the clusters themselves would be unrecognizable. The lower velocity cluster is dominated by the galaxy NGC4696, with NGC4709 having a similar role in the other cluster. The X rays presumably originate from the lower velocity cluster, as the source⁷ is coincident with NGC4696.

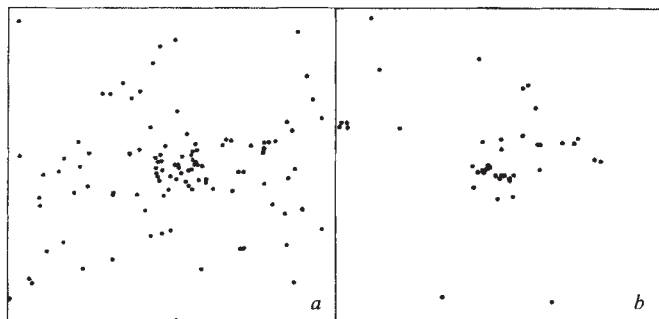


Fig. 2 *a*, The spatial distribution for galaxies with $V < 4,250 \text{ km s}^{-1}$ within the Schmidt field. *b*, The corresponding distribution for galaxies with $4,250 < V < 5,500 \text{ km s}^{-1}$.

The resolution of Centaurus I into two clusters leads us to believe that erroneous conclusions can be drawn about even well-studied clusters. In rich concentrations of galaxies there is bound to be some degree of superposition as it is the more enhanced regions that are picked out as clusters and groups. The existence of physically related, merging subunits will enhance the probability. Clearly contamination and substructure must be accurately assessed if derived cluster parameters are to be physically meaningful.

We thank M. J. Currie for galaxy magnitudes and D. J. King for help with the data reduction. J. R. L. acknowledges the receipt of an SRC studentship.

7. Mitchell, R. J., Charles, P. A., Culhane, J. L. & Davison, P. J. N. *Astrophys. J.* **200**, L5-L8 (1975).
8. Mitchell, R. J. & Culhane, J. L. *Mon. Not. R. astr. Soc.* **178**, 75P-80P (1977).
9. Gregory, S. A. & Thompson, L. A. *Astrophys. J.* **222**, 784-799 (1978).
10. Tarengi, M., Tift, W. G., Chincarini, G., Rood, H. J. & Thompson, L. A. *Astrophys. J.* **234**, 793-801 (1979).
11. Ulrich, M. *Astrophys. J.* **221**, 422-435 (1978).
12. White, S. D. M. *Mon. Not. R. astr. Soc.* **177**, 717-733 (1976).
13. Holmberg, E. B., Lauberts, A., Schuster, H. E. & West, R. M. *Astr. Astrophys. Suppl.* **34**, 285-340 (1978).

IR photometry and polarimetry of 2A0311-227

J. Bailey & J. H. Hough

Hatfield Polytechnic Observatory, Bayfordbury House,
Hertford SG13 8LD, UK

D. J. Axon

Astronomy Centre, University of Sussex, Falmer,
Brighton BN1 9QH, UK

and
Department of Physics and Astronomy, University College,
London WC1E 6BT, UK

The X-ray source 2A0311-227 has recently been identified with a star whose spectrum¹ and polarization properties² show it to be an AM Herculis type binary. These binaries are thought to consist of a synchronously rotating magnetic white dwarf, with a surface field of $\sim 10^8 \text{ G}$, accreting matter from a cooler companion star³. The highly circularly polarized radiation emitted by these systems is thought to be due to cyclotron radiation originating in an accretion column at one or both poles of the white dwarf. We describe here IR light curves of 2A0311-227 at 1.25 and 2.2 μm , together with circular polarimetry at 1.25 μm . The light curves show a deep narrow minimum which does not have an optical counterpart, and the 1.25 μm polarimetry shows circular polarization which is much lower than that in the optical.

The observations were obtained with the 3.8 m UK IR telescope at Mauna Kea. Photometry was obtained on 10 November 1979 in the *J* (1.25 μm) band and on 11 November in the *K* (2.2 μm) band, using the *f*/9 photometer. On 13 November the Hatfield IR polarimeter⁴ was used to obtain circular polarization observations at 1.25 μm . We have also obtained extensive optical photometry and polarimetry of 2A0311-227 using the 1.5 m telescope at the Cabezon Observatory, Tenerife, and the RGO Peoples photometer in its polarimetric mode⁵. These data will be described more fully elsewhere, but as well as IR results in Fig. 1 we present for comparison an optical light curve and circular polarization curve, obtained from observations in unfiltered light (extended S20 response) on 14-16 September 1979.

2A0311-227 shows a sharp linear polarization pulse repeating with its 81 min period, as first reported by Tapia². Using our polarization observations between 14 and 26 September 1979 and the epoch reported by Tapia² the following ephemeris has been derived.

$$T_{\text{lin pol max}} = \text{JD}_{\odot} 2444131.6751 + 0.0562666E \\ \pm 0.0000006$$

Photometry in the *J* band by Ward *et al.*⁶ showed a narrow minimum 0.8 mag deep. This feature can also be seen in our *J* light curve. Our *K* observations show that a narrow minimum is also present at this wavelength and is deeper than that at *J*. With our five minima (two at *J* and three at *K*) and the epoch of minimum observed by Ward *et al.*⁶ we can fit several possible periods, one of which is consistent with the

Received 27 November 1979; accepted 26 March 1980.

1. Turner, E. L., Aarseth, S. J., Gott III, J. R., Blanchard, N. T. & Mathieu, R. D. *Astrophys. J.* **228**, 684-695 (1979).
2. Lucey, J. R., Dickens, R. J. & Dawe, J. A. *Mon. Not. R. astr. Soc.* (submitted).
3. Sandage, A. *Astrophys. J.* **202**, 563-582 (1975).
4. Dawe, J. A., Dickens, R. J. & Peterson, B. A. *Mon. Not. R. astr. Soc.* **178**, 675-685 (1977).
5. Vidal, N. V. & Wickramasinghe, D. T. *Mon. Not. R. astr. Soc.* **180**, 305-308 (1977).
6. Mills, B. Y., Slee, O. B. & Hill, C. *Aust. J. Phys.* **13**, 679-699 (1960).

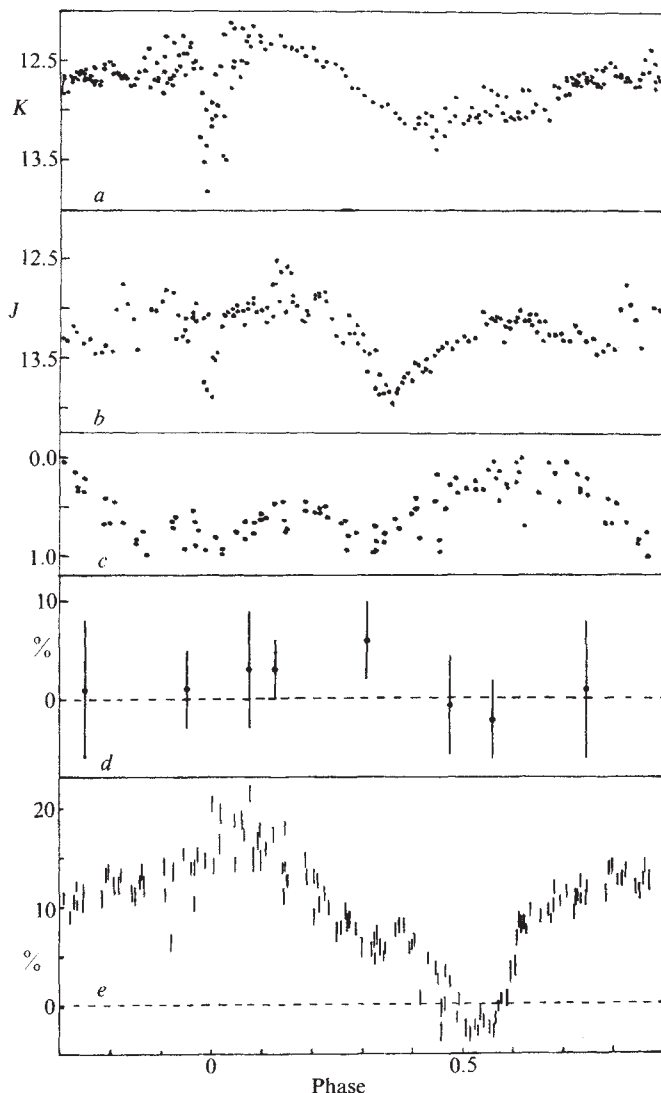


Fig. 1 *a*, *K* (2.2 μm) light curve obtained on 11 November 1979. Data from three cycles are plotted. Each point represents a 32-s integration. A typical point has a standard error of 0.08 mag. *b*, *J* (1.25 μm) light curve obtained on 10 November 1979. Data from two cycles are plotted. Each point represents a 32-s integration. A typical point has a standard error of 0.05 mag. *c* Broad band optical (0.4–0.85 μm) light curve from six cycles data obtained on 14–16 September 1979. The data are plotted on a magnitude scale with arbitrary zero point. *d* 1.25 μm circular polarization observations obtained on 13 November 1979. Each point represents a 10-min integration with standard error indicated. *e* Broad band optical (0.4–0.85 μm) circular polarization from observations of eight cycles on 14–16 and 26 September 1979.

polarization period. This corresponds to the following ephemeris:

$$T_{\text{IR min}} = \text{JD}_{\odot} 2443944.9518 + 0.0562660E \\ \pm 0.0000003$$

This ephemeris has been used to determine the phase for the observations plotted in Fig. 1. The narrow minimum seems to be slightly variable in phase and depth from cycle to cycle. The light curves also show flickering with amplitudes of up to 0.7 mag in *J*, and with a smaller amplitude in *K*. The flickering activity is greatest around the phase of the narrow minimum.

Observations by Tapia have been reported to show circular polarization always of the same sign⁷. However, our circular polarimetry shows a definite reversal of sign, lasting about 0.1 phase and reaching -2.5% in unfiltered light. Observations in

narrower bands show that the reversal is largest in blue light, reaching -7% in *B*. The linear polarization pulse occurs at phase 0.58 on the IR ephemeris, and corresponds to the change of sign in the circular polarization, as is usually the case in AM Her objects.

In contrast to the observed wavelength dependence of polarization in AM Herculis, which peaks at around 0.9 μm (refs 8,9) our optical data show a fall off in circular polarization towards the red. The peak circular polarizations observed in September 1979 were 15% at 0.44 μm , 29% at 0.55 μm , 20% at 0.68 μm and 9% at 0.83 μm . The observed 1.25 μm circular polarization follows this downward trend and is very much lower than that in the optical. The data are consistent with zero polarization, and averaged over the whole cycle give the value $3.0\% \pm 1.5\%$. Unfortunately no other IR circular polarimetry of AM Her type systems exists in the literature for comparison.

Watson, Mayo and King (personal communication) claim to detect a narrow minimum in the optical light curve at the same phase as the narrow IR minimum. They naturally interpret this as being an eclipse of the cyclotron emitting region by the companion star. However, this model is not supported by our data. The narrow optical minimum is not apparent in our light curve, and of greater significance, there is no minimum in the optical polarization curve at this phase, as an eclipse of the cyclotron region requires.

It is important to consider how our optical and IR observations may be reconciled. If the IR emission is due to cyclotron radiation then the optical and IR radiations must originate in different regions. We suggest two possibilities. (1) The IR and optical radiation originate at different heights in the same accretion column. (2) The IR emission is produced at a second pole which predominantly radiates in the IR. If the accretion rates of the two columns vary then it may also be possible to account for the difference between our optical curves and those reported by Watson *et al.*

If the narrow eclipse is due to the companion star then in either of the above cases severe limitations are imposed on the inclination of the system. Additionally both these explanations require some mechanism to depolarize the IR cyclotron radiation.

Alternatively some other mechanism could be responsible for the IR radiation, and dilution by this component could then explain the low polarization at long wavelengths. In this case, however, it is not clear what other mechanism could be consistent with the observed narrow eclipse as this imposes constraints on the maximum size of the emitting region. We have investigated the possibility that free-free emission from X-ray heated gas in the accretion stream is the source of the IR radiation. Assuming an emitting region with a radius $\sim 10^9 \text{ cm}$ the observed *K* luminosity $L_K \sim 5.7 \times 10^{16} (d/100)^2 \text{ erg s}^{-1}$ (where *d* is the distance in parsecs) yields a brightness temperature $T_B \sim 3 \times 10^6 (d/100)^2 \text{ K}$. The emission measure is so large that the gas will be optically thick to free-free radiation at all wavelengths considered.

With the data available we cannot offer a more complete explanation of our observations or the radial velocity data of Williams *et al.*¹⁰ and Schneider and Young (personal communication) which depends on the detailed geometry of the accretion scheme. Clearly a detailed understanding of 2A0311–227 requires further coordinated optical and IR measurements.

We thank the UKIRT staff for their assistance and D. J. Adams for help with the UKIRT observations. The RGO is thanked for making available the Peoples photometer and the excellent technical support of M. Waite for our Tenerife observations. M. J. Ward and D. A. Allen provided results in advance of publication. We thank A. King for stimulating discussions and S. Tapia for helpful comments. PATT is acknowledged for allocations of time at UKIRT and Tenerife. D.J.A. is supported by an SRC fellowship.

Received 27 December 1979; accepted 27 March 1980.

1. Griffiths, R. E., Ward, M. J., Blades, J. C., Wilson, A. S. & Chaisson, L. *Astrophys. J.* **232**, L27 (1979).
2. Tapia, S. *IAU Circ. No.* 3327 (1979).
3. Channugam, G. & Wagner, R. L. *Astrophys. J. Lett.* **213**, L13 (1977).
4. Cox, L. J., Hough, J. H. & McCall, A. *Mon. Not. R. astr. Soc.* **185**, 199 (1978).
5. Bailey, J. thesis, Univ. Sussex (1978).
6. Ward, M. J., Allen, D. A., Smith, M. G. & Wright, A. E. *IAU Circ. No.* 3335 (1979).
7. Bond, H. E., Channugam, G. & Grauer, A. D. *Astrophys. J.* **234**, L113 (1979).
8. Michalsky, J. J., Stokes, G. M. & Stokes, R. A. *Astrophys. J. Lett.* **216**, L35 (1977).
9. Priedhorsky, W. C., Krzeminski, W. & Tapia, S. *Astrophys. J.* **229**, 542 (1978).
10. Williams, G. *et al.* *Nature* **281**, 48–49 (1979).

An upper limit to the global SO₂ abundance on Io

P. S. Butterworth*, J. Caldwell†, V. Moore†§, T. Owen†, A. R. Rivolo† & A. L. Lane‡

*Department of Astronomy, University of Leicester, Leicester LE1 7RH, UK

†Department of Earth and Space Sciences, State University of New York, Stony Brook, New York 11794

‡Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, California 91103

Gaseous sulphur dioxide has been detected on Io by the Voyager 1 IRIS experiment¹. It was associated with a major volcanic plume, and it was observable only because the local underlying surface was warm enough (>210 K) to be detected near 7 μ m, and thereby reveal the gas in absorption. The estimated column abundance of SO₂ was ~0.2 cm atm. At most other locations on Io, the surface was not warm enough to be detected at this wavelength (S. Kumar, personal communication). A refined Io model for an SO₂ atmosphere in equilibrium with the solid at local surface temperatures predicts a global average column abundance of SO₂ of 0.032 cm atm (J. Pearl, personal communication). Furthermore Smythe *et al.*² and Fanale *et al.*³ have shown the similarity between ground-based observations of the reflectivity of Io from 1.0 to 4.2 μ m and laboratory observations of SO₂ frost. We demonstrate here that the near UV reflectivity of Io, measured from Earth-orbit, demands that the hemispherical average column abundance of SO₂ at the time of our observation was very much less than the local abundance quoted above. Our limit is also less than the average abundance the model would predict, although we do find the observations to be consistent with the presence of condensed SO₂ on Io's surface. A qualitative reconciliation of these results is presented.

Io was observed by the long-wavelength spectrograph of the IUE on 21 April 1979. The best Io spectrum is shown in Fig. 1, together with the solar spectrum of Broadfoot⁴. The sub-Earth longitude on Io increased from 80 to 90° during the period in which our best spectrum was obtained. The spectral resolution for the Io data is 8 Å; the solar data have been convolved with a rectangular 3 Å bandpass. This treatment of the solar spectrum was chosen to provide a slightly better resolution than achieved by the IUE so there would be no doubt about the reality of any non-solar absorptions found in the spectrum of Io. The exposure time on Io was 40 min, but the satellite was probably outside the spectrograph aperture for a significant fraction of this time because of guiding uncertainties. The small motions of Io within and just outside the aperture reduce the signal uniformly at all wavelengths and slightly degrade the spectral resolution. However, neither effect seriously impairs these data.

Figure 1 indicates that all the features in the spectrum of Io from 2,900 Å to 3,100 Å are of solar rather than ionian origin.

We conservatively estimate that no absorption features occur on Io at the 20% level. Also shown in Fig. 1 are the locations of the strong electronic bands of SO₂ in this region⁵. We note that SO₂ has much stronger bands near 2,100 Å, but both the instrument's insensitivity and the intrinsic faintness of Io at shorter wavelengths render the 3,000 Å region most suitable for our present study.

Adopting a typical valley-to-peak absorption coefficient of 10 (cm atm)⁻¹ for SO₂ in this spectral region, assuming the upper limit of 20% for any absorption feature, and using an air-mass factor of 3, we calculate an upper limit of 0.008 cm atm of SO₂. Because we have been very conservative in choosing the upper limit for undetected percentage absorption, and because there are more than 10 candidate bands to be sought, we believe our upper limit on the abundance is quite firm, and must be reconciled with the IR findings.

As a further test, we have produced the ratio spectra shown in Fig. 2. Previous experience with IUE spectra has shown that slight differences between the wavelength scales and slit functions of the IUE and the convolved solar spectrum produce high but irrelevant noise levels. The ratio spectra of Vesta and Io illustrated in Fig. 2 were obtained by using the same treatment of the solar spectrum that was used in Fig. 1. Both of these ratios are tied to OAO 2 observations of the geometric albedos of these objects⁶.

The ratio spectra are indeed noisy, but their basic similarity is easily seen. Over the wavelength range 2,500–3,000 Å, there is no indication of any absorption in the spectrum of Io that is not also present on Vesta. Despite the noise in these spectra, we estimate that any broad deviation larger than 30% of the signal would be detectable. Adopting 15 (cm atm)⁻¹ as the average total absorption coefficient in this region and again assuming an air-mass of 3, we again find an upper limit of 0.008 cm atm for the average SO₂ column abundance over the visible hemisphere.

The ratio spectra of Io and Vesta in Fig. 2 diverge rapidly at longer wavelengths, with Io becoming brighter than Vesta as we approach 3,200 Å. Beyond 3,050 Å, the IUE calibration is less certain than at shorter wavelengths⁷. Nevertheless, we have found this same trend in three separate Io spectra, while it is absent from all other spectra (several asteroids, Neptune,

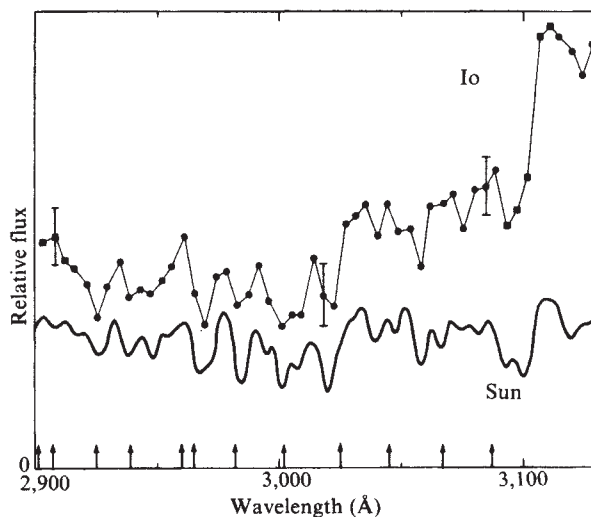


Fig. 1 UV spectra of Io, from IUE, and of the convolved Sun, from ref. 4. The arrows along the abscissa give the locations of strong SO₂ bands⁵, the absence of which in the spectrum of Io leads to the upper limit given in the text. The error bars on the Io data are derived from the background measured on the spectral image, both above and below the spectrum. Individual Io data points are plotted as filled circles. Instrumental reseau, indicated by the squares near 2,900 and 3,100 Å, have no discernible effect on this spectrum. See, however, Fig. 2.

§Permanent address: Department of Physics and Astronomy, University College London, London WC1E 6BT, UK

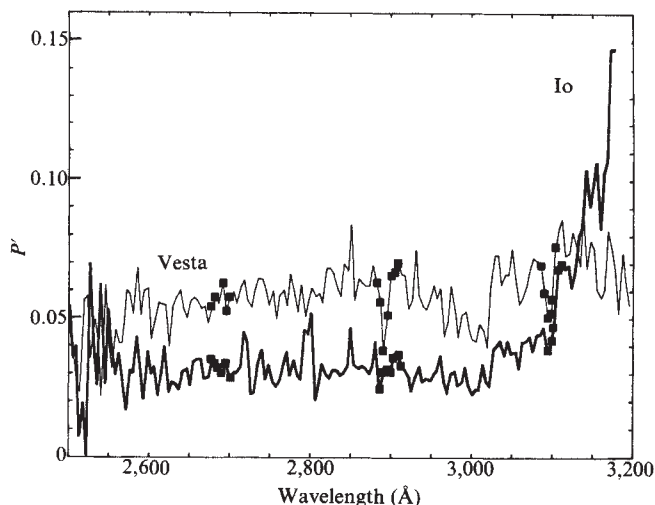


Fig. 2 Ratio spectra of Vesta and Io normalized to geomatrix albedos at 2,590 Å determined by OAO⁶. Points affected by reseaux are marked with black squares. The ordinate gives the reflectivity P' as this normalized albedo. Note that the signal is lost in the noise below 2,550 Å, and that the reflectance of Io crosses that of Vesta at 3,150 Å. Both spectra become progressively less reliable at wavelengths longer than 3,050 Å, but the relative upward trend in the Io reflectance is real.

Titan, Ganymede, Venus, Jupiter, Callisto) we have checked. This rapid rise is also consistent with ground-based observations⁸, which unfortunately begin to be uncertain at just this same wavelength. We note that the reflectance of SO₂ frost deposited at 130 K is less than 3% from 2,600 to 3,000 Å, rising very steeply at slightly longer wavelengths: 47% at 3,250 Å, 68% at 3,500 Å and so on⁹. The present UV observations are thus consistent with the conclusions from IR spectra of Io^{2,3} that SO₂ frost is present on the satellite's surface, but they are by no means diagnostic. For example, they are also consistent with deposits of Na₂S and K₂S (ref. 10).

There remains the apparent discrepancy between our upper limit for gaseous SO₂ and the result based on the detection by Pearl *et al.*¹. One possibility is that the UV absorption coefficient could vary with temperature. While such a variation may occur from band to band, we would expect the total absorption to remain constant. In this circumstance, the calculated upper limit would be based on the stronger bands, and it would probably decrease further. (We thank E. S. Barker for calling our attention to this possibility.)

Transience on Io would also explain the discrepancy. The Voyager observations were made in early March, 1979, whereas the IUE data were obtained in mid-April. It would not be surprising for an atmospheric constituent expelled from volcanoes to be time-variable. This possibility is further enhanced by the Voyager 2 finding that one major active volcanic vent from the Voyager 1 epoch had become dormant by July 1979 (ref. 11).

What seems the most probable explanation, however, is that the SO₂ 'atmosphere' on Io is patchy, being confined to regions over solid deposits and sources of volcanic venting. We hypothesize that the SO₂ gas condenses locally on the surface, either as a frost or as snow, producing the white patches visible on Voyager pictures¹¹. Thus the fact that our upper limit is a factor of 4 times smaller than the global average predicted by Pearl's model is simply an indication that the surface of Io is not completely covered by SO₂ frost, which is certainly consistent with the satellite's appearance. This hypothesis is also consistent with refs 1, 2 and 3, and with the upper limit on the surface pressure set by Pioneer 10 occultation experiments¹². Our upper limit, which is a global

average, does not in any case contradict the Voyager-IRIS result for gaseous SO₂, which applies only to a very small fraction (<1%) of the surface of Io. Our interpretation leaves open the possibility for small variations in global average absorption of both solid and gaseous SO₂ with time and with the sub-Earth longitude on Io, but such variations may well remain below the detection threshold of the IUE.

We thank P. Solomon for use of his graphics facilities and A. T. Sinclair for providing the ephemeris of Io. We thank J. Pearl and F. Fanale for unpublished information and B. Buratti for help with data analysis. This research was supported in part by NASA grants NSG 5250 and NGR 33-015-141 at Stony Brook. P.S.B. and V.M. were supported by the SRC.

Received 5 October 1979; accepted 15 March 1980.

1. Pearl, J. *et al.* *Nature* **280**, 755 (1979).
2. Smythe, W. D., Nelson, R. M. & Nash, D. B. *Nature* **280**, 766 (1979).
3. Fanale, F. P., Brown, R. H., Cruikshank, D. P. & Clark, R. N. *Nature* **280**, 761 (1979).
4. Broadfoot, A. L. *Astrophys. J.* **173**, 681 (1972).
5. Warneck, P., Marmo, F. F. & Sullivan, J. O. *J. chem. Phys.* **40**, 1132 (1964).
6. Caldwell, J. *Icarus* **25**, 384 (1975).
7. Bohlin, R. C. & Snijders, M. A. J. *IUE Calibration Memo VI* (September 1978).
8. Nelson, R. M. & Hapke, B. W. *Icarus* **36**, 304 (1978).
9. Nash, D. B., Fanale, F. P. & Nelson, R. M. *Bull. Am. astr. Soc.* **11**, 597 (1979).
10. Nash, D. B. & Nelson, R. M. *Nature* **280**, 763 (1979).
11. Smith, B. A. *et al.* *Science* **204**, 951, 206, 927 (1979).
12. Kliore, A. *et al.* *Icarus* **24**, 407 (1975).

The terminal Eocene event: formation of a ring system around the Earth?

John A. O'Keefe

Laboratory for Astronomy and Solar Physics, Goddard Space Flight Center, Greenbelt, Maryland 20771

The most profound climatic event of the Tertiary was the terminal Eocene event at -34 Myr (ref. 1). Botanical data indicate that winters became much more severe, while summer temperatures were little affected. An explanation in terms of a change in the space direction of the Earth's axis is not dynamically acceptable. On the other hand, an ecological disaster of some kind apparently struck the Radiolaria at this time²; the latter event is accurately correlated (to a few tens of thousands of years) with the formation of the greatest known tektite strewn field, the so-called North American strewn field, which has recently been shown to extend at least half-way around the Earth³. It is suggested here that tektites and microtektites which accompanied this fall, but missed the Earth, organized themselves into a ring system like that of Saturn. The shadow of the rings fell on the winter hemisphere and so produced the observed cooling. The ring lasted between one and several million years.

Wolfe, reviewing the botanical data, found a sudden change in the relative abundance of forest plants at the end of the Eocene, at about -34 Myr which, he suggests, implies a decrease of the winter temperature by ~20°C, with little or no corresponding change of summer temperature. Wolfe attributes the event to a change in the obliquity of the ecliptic from 5°-10° before the event, to 25°-30° after it. Other calculations^{4,5} have shown, however, that although changes as large as this have taken place in the obliquity, they have occurred on a time scale of billions of years; they are much too slow to control Eocene climatic changes.

Urey⁶ suggested that this and other Cenozoic extinctions may have been associated with tektite strewn fields (showers of tektites). The largest known strewn field is dated at -34 Myr,

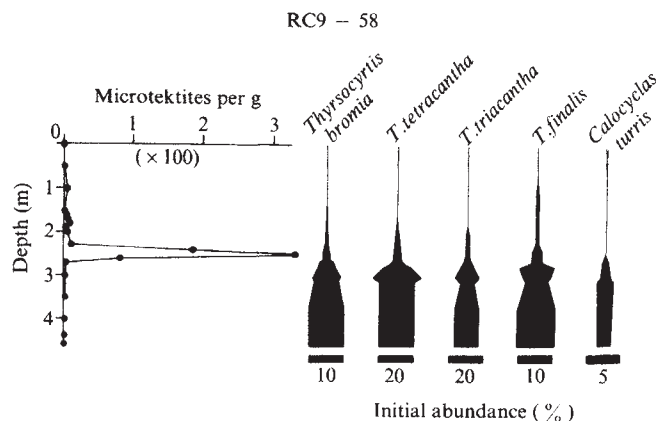


Fig. 1 Microtektites and radiolarian extinctions in a Caribbean deep-sea core. The top of the core is near the end of the Eocene. One metre of core depth corresponds to 0.1–0.2 Myr. The bar graphs display the varying abundances of five species of Radiolaria. The initial abundance of each species, relative to all Radiolaria, is marked. Burrowing animals have disturbed the sediment slightly. From Glass and Zwart² by permission.

close to the time of the terminal Eocene event. It has been called the North American strewn field; but Glass *et al.*³ found that it extends from the Caribbean through the central Pacific to the Indian Ocean, and that its total mass is $1\text{--}10 \times 10^9$ tons. Glass and Zwart² have shown, from ocean-bottom cores, that five abundant species of Radiolaria, comprising two-thirds of the total radiolarian population, disappear or nearly disappear at the level at which the microtektites appear. Figure 1 compares the variations in the abundance of these species with depth in core RC 9-58, from the Caribbean. Glass estimates the sedimentation rate as $5\text{--}10 \text{ Myr}^{-1}$ in this core; he considers that the agreement between the date of the extinctions seen in Fig. 1 and that of the tektite strewn field is within a few tens of thousands of years (personal communication).

If there is a connection between tektites and climatic change, then it probably results from the screening of sunlight. The mass of even the largest strewn fields is insufficient to affect directly the dynamics of the Earth's atmosphere, not to mention the hydrosphere or the lithosphere. The screening probably also took place in space: dust in the troposphere washes out in a few weeks; in the stratosphere it falls out in a few years at most; but the evidence is for a climatic change that lasted one to several million years: dust in space can persist for that time.

The tektites are assumed here to be of cosmic origin⁷; a terrestrial origin implies the almost instantaneous production of homogeneous, non-porous, water-free glass in pieces up to several decimetres in length, out of common rocks or soil. As far as is known, glass cannot be made in this way.

Assuming a cosmic origin, it would be expected that, in addition to the tektites which reach the Earth, there would be a larger mass which would miss the Earth and be trapped in geocentric orbit. Glass *et al.*³ estimate $1\text{--}10 \times 10^9$ ton for the mass of the North American strewn field; let us adopt 25×10^9 ton, or 10 km^3 , for the matter trapped in orbit. Typical diameter sizes for microtektites found in the oceanic cores are of the order of $100 \mu\text{m}$; this would imply about 2×10^{22} particles with a total area of cross-section of $1.6 \times 10^8 \text{ km}^2$.

Initially the orbits of the trapped particles would form a cloud about the Earth, densest at the inner surface. The cloud would rapidly collapse to a thin ring in the plane of the Earth's equator⁸; the time scale t_1 for the collapse is given by Brahic⁹ as

$$t_1 = \frac{T}{N} \left(\frac{R}{r} \right)^2$$

where T is the orbital period; N the number of particles; R a representative dimension (radius of the cloud); and r the particle radius. Then t_1 turns out to be about half a day, so that the reduction to a thin ring will be complete after a year.

An inner radius of 1.5 Earth radii is assumed, above ordinary atmospheric drag, and an outer boundary of 2.5 Earth radii is assumed, inside the Roche limit.

As the ring would lie exactly in the equatorial plane (as a result of collisional damping of north-south motion) it follows that it would cut off sunlight during the winter months, but not the summer months, as shown in Fig. 2.

Quantitatively, the total area of cross-section of the particles would be 30% of the total surface area of the rings; hence the vertical optical depth τ would be 0.30. The Sun's rays would, however, make an angle with the ring equal to the Sun's declination δ , which has a range of $\pm 23.5^\circ$. If the particles were perfect absorbers, the fraction f of the light which would penetrate the ring would be given by

$$f = \exp(-\tau \csc \delta)$$

and f would vary from 0–0.47. Glass and Zwart¹⁰ report three kinds of particles in the North American strewn field: transparent, opaque white, and black. The data are thus consistent with a reduction of the light by a factor of 0.75, as would be required to produce a 20° fall in temperature (grey body calculation).

Estimation of the ring lifetime requires consideration of eight effects:

(1) Gravitational clumping is not expected inside the Roche limit. (2) Tidal effects due to the attraction of the Moon will be about twice as large as the interesting but minor effects of Mimas on the rings of Saturn¹¹. (3) Collisional spreading in

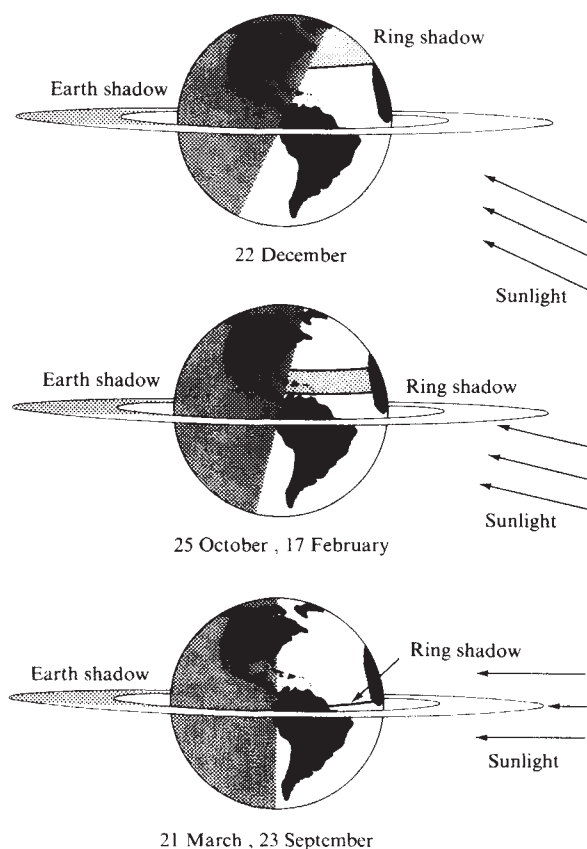


Fig. 2 The terrestrial ring, showing the movement of the shadow with the seasons. Continents shown in their present position. Dates correspond to solar declinations of 0° , 12° , 23.5° . Note that the tropical zone experiences a secondary maximum of illumination around winter solstice.

the ring plane has a time scale of the order of 10^{18} yr (ref. 9); it may be neglected. (4) Atmospheric drag is eliminated because the gas molecules are mopped up by the ring¹². (5) Micrometeorite erosion at the lunar rate ($0.2\text{--}0.8\text{ mm Myr}^{-1}$) would severely limit the lifetimes of the particles; but most of the erosion in the lunar case is due to craters $>400\text{ }\mu\text{m}$ in diameter¹³; the small size of the ring particles has a large effect in limiting damage. (6) Rotational bursting due to windmill effects of radiation pressure¹⁴ will be damped by collisions. (7) Thrust effects due to photon recoil¹⁵ will also be damped by collisions. (8) The Poynting–Robertson effect would cancel out around a geocentric orbit except for the fact that part of the orbit lies in the Earth's shadow. The net effect is an outward motion (if the sense of the orbit motion is direct) which probably limits the lifetime of the ring to a few million years.

I thank A. F. Cook, B. P. Glass and M. B. Swincki for their unpublished results and D. E. Smith, J. A. Wolfe and D. P. Rubincam for helpful discussions.

Note added in proof: Ruderman and Truran¹⁶ have similarly suggested that dust ejected from the Moon may cause sudden climatic changes on the Earth.

Received 2 January; accepted 8 April 1980.

1. Wolfe, J. A. *Am. Scient.* **66**, 694–703 (1978).
2. Glass, B. P. & Zwart, M. J. in *Proc. Symp. Stratigraphic Micropaleontology of the Atlantic Basin and Borderlands* (ed. Swain, F. W.) 553–567 (Elsevier, New York, 1977).
3. Glass, B. P., Swincki, M. B. & Zwart, P. A. *Lunar planet. Sci.* **10**, 434–436 (1979).
4. Darwin, G. H. *Scientific Papers Vol. 2*, 208–302 (Cambridge University Press, 1908).
5. Brosche, P. & Sündermann, J. (eds) *Tidal Friction and the Earth's Rotation* (Springer, New York, 1978).
6. Urey, H. C. *Nature* **242**, 32–33 (1973).
7. O'Keefe, J. A. *Tektites and their Origin*, 180–187 (Elsevier, New York, 1976).
8. Jeffreys, H. *Mon. Not. R. astr. Soc.* **107**, 263–267 (1947).
9. Brahic, A. *Astr. Astrophys.* **54**, 895–907 (1977).
10. Glass, B. P. & Zwart, M. J. *Bull. geol. Soc. Am.* **90**, 595–602 (1979).
11. Goldreich, P. & Tremaine, S. *Icarus* **34**, 240–253 (1978).
12. Opp, A. G. *Science* **207**, 401–403 (1980).
13. Comstock, G. M. *Lunar Science VII*, 202–204 (Lunar Science Institute, Houston, 1977).
14. Paddack, S. J. *J. geophys. Res.* **74**, 4379–4381 (1969).
15. Jacchia, L. in *The Moon, Meteorites and Comets* (eds Middlehurst, B. M. & Kuiper, G. P.) 783–784 (University of Chicago Press, Chicago, 1963).
16. Ruderman, M. & Truran, J. W. *Nature* **284**, 328–329 (1980).

Bivalent spectator oxo bonds in metathesis and epoxidation alkenes

Anthony K. Rappé & William A. Goddard III

Arthur Amos Noyes Laboratory of Chemical Physics, California Institute of Technology, Pasadena, California 91125

Several studies suggest the importance of oxygen in the reactive intermediates for alkene metathesis,



and of additional metal=oxo bonds in epoxidations,

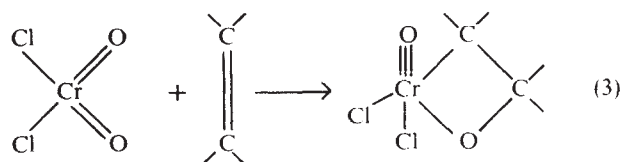


For example: (1) catalytic amounts of oxygen were found to be essential for the formation of a stable metathesis catalyst in the $\text{WCl}_6/\text{C}_2\text{H}_5\text{AlCl}_2$ system¹; (2) alkene epoxidation (and general oxidation) reactants such as CrCl_2O_2 , OsO_4 , and SeO_2 all have at least two oxygen ligands²; and (3) recently Osborn (personal communication) has described an $\text{Mo}=\text{oxo}$ carbene as an effective metathesis catalyst. From extensive *ab initio* theoretical studies of Cr and Mo complexes, we show here that the presence of spectator metal=oxo bonds drives the formation of metallocycles, thereby playing a crucial role in the chemistry.

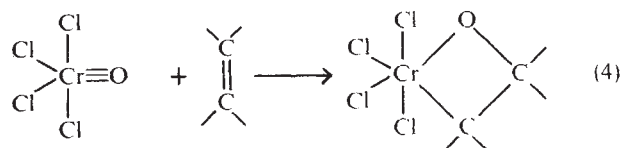
The special role of the oxo group is illustrated by comparing the oxo bond energy of $\text{Cl}_4\text{Cr}=\text{O}$ (343 kJ mol^{-1}) with the oxo bond energy of $\text{Cl}_2\text{Cr}=\text{O}$ (213 kJ mol^{-1}). (Calculated numbers are based on *ab initio* calculations using valence double zeta basis sets and including electron correlation (generalized valence bond plus configuration

interaction as in ref. 3). Estimated vibrational frequencies were used to calculate ΔH (0 K) from the theoretical energy differences. Standard approaches⁴ were used to estimate ΔS and the correction of ΔH to 25°C .) The origin of this difference is that in $\text{Cl}_4\text{Cr}=\text{O}$ there are two $d\pi$ orbitals on the Cr available for forming π bonds to the oxygen, leading to a triple bond (analogous to $\text{C}\equiv\text{O}$). However, in $\text{Cl}_2\text{Cr}=\text{O}$ the corresponding d orbitals are each used in the π bond for different oxo bonds, leading to two double bonds (analogous to $\text{O}=\text{C}=\text{O}$ or a ketone, $\text{R}_2\text{C}=\text{O}$). Thus the metal=oxo bond is dual-valent, existing in two alternative forms depending on the other ligands. The orbital character of the oxo double bond is as in a ketone⁵ with a singly-occupied $\text{O}p\sigma$ orbital paired with the singly-occupied metal $d\sigma$ orbital, a singly-occupied $\text{O}p\pi_x$ orbital paired with a singly-occupied $d\pi_x$ orbital and the remaining doubly-occupied $\text{O}p\pi_y$ pair nonbonding. For the oxo triple bond, the oxygen orbitals are oriented so as to have two π bonds with the $p\sigma$ pair of the oxygen coordinated with an empty metal $d\sigma$ orbital. In both cases the bonding is basically covalent with a small charge on the oxygen.

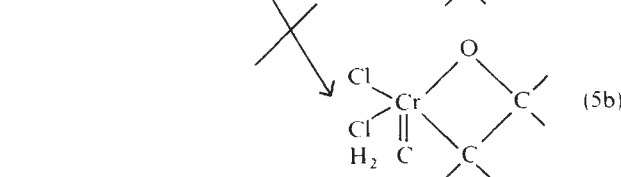
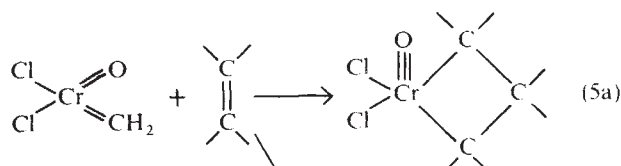
Addition of an alkene to Cl_2CrO_2 to form the metallocycle oxetane from one oxo bond allows the spectator oxo bond to convert from a double bond to a triple bond, thereby providing a 130 kJ mol^{-1} push towards formation of the metallocycle (calculated $\Delta H = -113\text{ kJ mol}^{-1}$, $\Delta G(25^\circ\text{C}) = -63\text{ kJ mol}^{-1}$),



For a single oxo bond, this process is doubly bad; here there is no spectator oxo bond to drive the metallocyclization and one must also add across an oxo triple bond (calculated $\Delta G(25^\circ\text{C}) = +300\text{ kJ mol}^{-1}$),



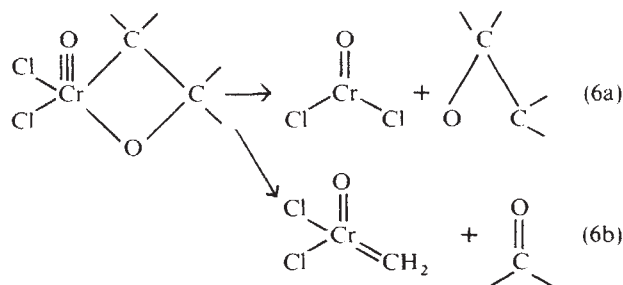
An oxo-carbene behaves similarly,



with metallocyclobutane formation favoured by conversion of the spectator oxo bond to a triple bond (calculated $\Delta G(25^\circ\text{C}) = -42\text{ kJ mol}^{-1}$), whereas metallocyclohexane formation has no such driving force and is significantly uphill (calculated $\Delta G(25^\circ\text{C}) = +50\text{ kJ mol}^{-1}$).

The same effects are found from calculations on Mo [$\Delta G = -105\text{ kJ mol}^{-1}$ for reaction (3), $\Delta G = -84\text{ kJ mol}^{-1}$ for reaction (5a)] and should occur also for W. The essential difference between Cr and Mo (or W) is in the decomposition products of the metallocycles and arises from the change in the

strength of metal carbon or metal oxygen bonds. For $\text{Cl}_4\text{Mo}\equiv\text{O}$ we find an oxo bond energy of 456 kJ mol^{-1} , whereas for $\text{Cl}_4\text{Cr}\equiv\text{O}$ we find 343 kJ mol^{-1} . For example, for Cr metalocycle oxetane we find that reductive elimination dominates [reaction (6a)], whereas for Mo, metal carbene formation is competitive [reaction (6b)]. Similarly, Cr metalocyclobutane reductively eliminates propane, whereas for Mo, the metathesis product is formed,



We suggest that this dual-valent character of the spectator metal-oxo bond may be generally useful for designing catalysts or substrates to promote reactions involving metalocycle formation.

This research was partially supported by the US Department of Energy Research and Development Administration grant EX-76-G-03-1305. However, any opinions, findings, conclusions, or recommendations expressed herein are those of the authors and do not necessarily reflect the view of the DOE.

Received 9 January; accepted 27 March 1980.

1. Mocella, M. T., Rovner, R. & Muetterties, E. L. *J. Am. chem. Soc.* **98**, 4689 (1976).
2. Sharpless, K. B., Teranishi, A. Y. & Bäckvall, J. E. *J. Am. chem. Soc.* **99**, 3120 (1977).
3. Harding, L. B. & Goddard III, W. A. *J. Am. chem. Soc.* **99**, 4520 (1977).
4. Benson, S. W. *Thermochemical Kinetics* (Wiley, New York, 1976).
5. Harding, L. B. & Goddard III, W. A. *J. Am. chem. Soc.* **97**, 6293 (1975).

Observations of nitrous acid in an urban atmosphere by differential optical absorption

U. Platt*, D. Perner*, G. W. Harris†, A. M. Winer† & J. N. Pitts Jr†

*Kernforschungsanlage Jülich GmbH, Institut für Chemie 3: Atmosphärische Chemie, Postfach 1913, D-5170 Jülich, FRG

†Statewide Air Pollution Research Center and Department of Chemistry, University of California, Riverside, California 92521

In the past decade, increasing attention has been paid to the role of nitrous acid (HONO) in the polluted troposphere, primarily as an initiator of photochemical air pollution¹ through its photodissociation by solar UV radiation into hydroxyl radicals and NO. The OH radical may subsequently attack organics starting a chain photooxidation which leads to the production of O_3 , peroxyacetyl nitrate and many other secondary pollutants. Measurements of the concentration of HONO, especially in the early morning, are needed to establish the initial conditions to be used in computer kinetic models of photochemical oxidant formation²⁻⁴. From laboratory studies, it has been suggested that nitrous acid may also be a precursor to the possible formation of nitrosamines by reaction with simple secondary and tertiary amines in urban air⁵, or *in vivo* following inhalation. We report here a series of observations, at Riverside and Claremont, California, of the gradual buildup of HONO during the night and its rapid decay after sunrise. These, we believe, represent the first unequivocal measurements of HONO reported for a major urban air basin impacted by photochemical air pollution.

Nitrous acid was detected and monitored using the long path differential UV absorption technique described by Platt *et al.*⁶. The optical system in the present experiments used a 0.3 m McPherson spectrograph with a 600 grooves mm^{-1} grating blazed at 500 nm (giving a dispersion of 5.3 nm per mm). A spectral region 38 nm wide was scanned by exit slits etched in a thin metal disk, which rotated in the focal plane of the instrument.

The light signals were received by a photomultiplier (EMI 9656Q) and averaged by a DEC PDP11 MINC microcomputer. The scan rate of 105 Hz provided by the rotating disk was fast enough to eliminate unwanted signal variations. Absorption lines with optical densities as low as 10^{-4} (base 10) could be detected. We used the HONO absorption lines at 354.1 and 368.1 nm which have differential absorption cross-sections (D.P., unpublished data) of 4.7×10^{-19} and $3.8 \times 10^{-19} \text{ cm}^2$ respectively. Before measuring the optical density of atmospheric HONO bands, nearby NO_2 bands were eliminated by the subtraction of a suitably weighted NO_2 reference spectrum. For the 354.1 nm band the minimum detectable optical density corresponded to 0.28 p.p.b. (parts per 10^9) HONO using a 970 m lightpath. The light source was a 500-W Xe high pressure lamp (Hanovia 959C).

The night-time buildup of HONO was measured during August and September 1979 (see Table 1) in the Los Angeles air basin at Riverside (a downwind receptor site) and at Claremont (a midbasin site). Figure 1 shows the development of HONO bands during the night of 4-5 August and their decay after sunrise. The observed bands agree well in shape and position with the reference spectrum shown. Additional broad absorption features at ~ 352 and 360 nm are present

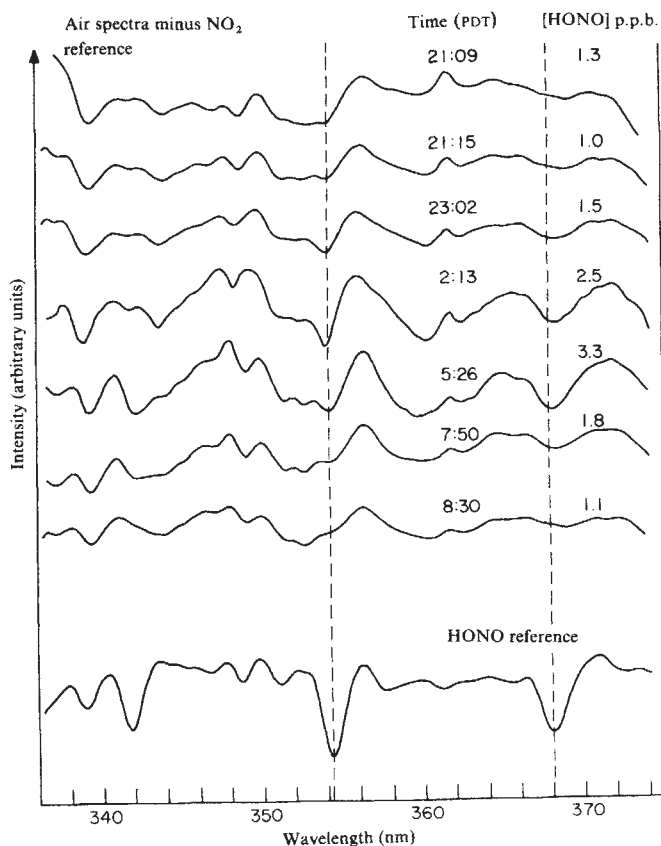


Fig. 1 Air spectra from which absorptions due to NO_2 have been subtracted. 4-5 August 1979, Riverside, California. The lowest trace is a HONO reference spectrum.

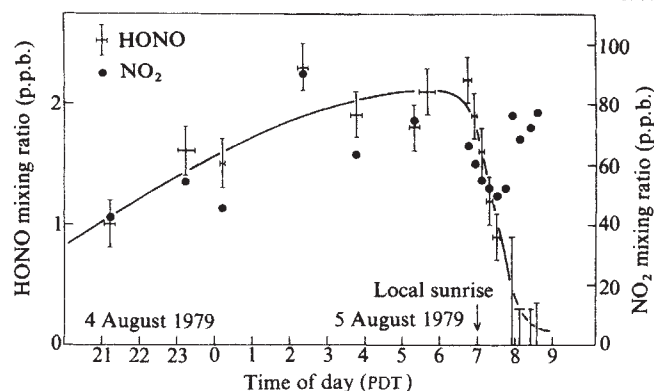
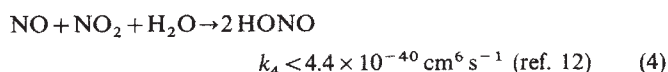
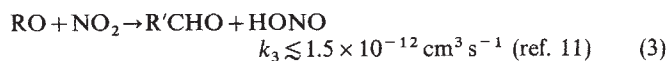
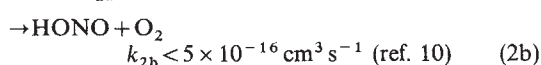
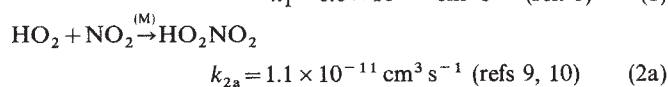
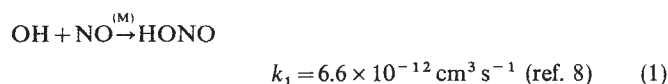


Fig. 2 Concentration-time profiles of HONO and NO₂. 4-5 August 1979, Riverside, California.

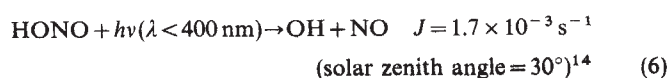
and are possibly due to hydrocarbons but are as yet unassigned; the narrow absorption near 339 nm is partly due to CH₂O. At our resolution, deconvolution of this band and the 342 nm HONO band present in the reference spectrum is difficult, leading to distortion and masking of the HONO band. This is particularly evident in the spectrum taken at 2:13 in Fig. 1. Figure 2 shows the HONO and NO₂ concentrations measured on the night of 4-5 August, and Fig. 3 shows the decay of HONO near sunrise during that and three other mornings. Table 1 gives the maximum HONO concentrations observed and the maximum rates of decay after local sunrise. These rates varied between ~0.3 and 2.6 p.p.b.h⁻¹, being slower and occurring longer after local sunrise, on hazy mornings as expected.

Nitrous acid was not present above our detection limit during daylight hours and hence we will confine our discussion to the night-time and early morning chemistry of HONO. Our detection limit (~0.28 p.p.b.) was relatively high compared to that reported for a study at Jülich, West Germany⁷, as we had to use a shorter pathlength (970 m at Riverside and 750 m at Claremont) because of the poor optical transmittance of the polluted atmosphere at our observation sites. In the earlier study⁷, despite a longer pathlength, HONO could only be detected on a few occasions in the morning before sunrise while during the remainder of the day concentrations were close to or below the detection limit.

Reactions which are known or have been suggested to produce HONO in the troposphere are



and of these, only reactions (4) and (5) can proceed at night in the absence of photolytic radical sources. During daylight HONO is destroyed rapidly by photolysis



but at night it should accumulate at a rate reflecting the source strength. From the observed morning concentrations of 1.5 to 4.1 p.p.b., mean source strengths of 0.13 to 0.34 p.p.b.h⁻¹ (~1–3 × 10⁶ cm⁻³s⁻¹) can be estimated.

Kaiser and Wu¹² found that the rate of the homogeneous process reaction (4) is very slow ($k_4 < 4.4 \times 10^{-40} \text{ cm}^6 \text{ s}^{-1}$) and for reactant concentrations typical of the night-time Riverside atmosphere (NO₂ ~ 80 p.p.b., NO < 1 p.p.b., H₂O ~ 2 vol %) the homogeneous reaction is negligible in a well-mixed air mass. The very low levels of NO also rule out reaction (5) as a major HONO source.

In motor vehicle exhaust systems only a few p.p.t. (part per 10¹²) of HONO can be formed by the homogeneous reaction (4) during the 0.2 s residence time of the exhaust gases in the silencer of an average car with its engine running at 3000 r.p.m. (NO_x concentration ~ 100 p.p.m. and NO/NO₂ ~ 10). Moreover, the homogeneous formation of HONO in plumes emanating from mobile or stationary sources, even though they provide much longer residence times, can also be neglected because the rate of HONO production decreases with the third power of the reaction partner concentrations and dilution by ambient air is rapid. Thus, if reaction (4) is important in the atmosphere, it must proceed heterogeneously, as suggested by Kaiser and Wu¹² who

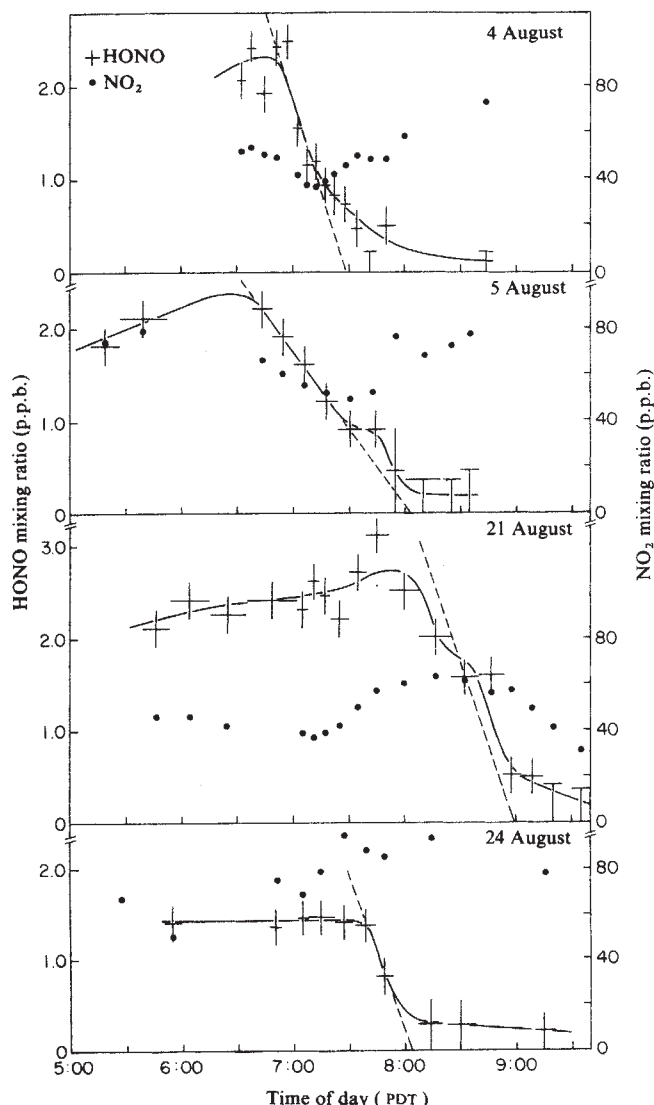


Fig. 3 Measurements of HONO and NO₂ performed near sunrise on four days during August 1979 at Riverside, California.

Table 1 Data for HONO observation at Riverside and Claremont, California

1979 date	HONO maximum (p.p.b.)	Max HONO decay rate (p.p.b. h ⁻¹)	HONO half-time* (PDT)	Sunrise† (PDT)
8/2	4.1	2.6	7:10	6:11
8/4	2.5	2.4	7:10	6:13
8/5	2.2	1.7	7:15	6:14
8/21	3.1	1.5	8:45	6:26
8/24	1.5	2.1	7:45	6:28
8/29‡	1.3	~0.3	~9:30	6:31
8/31‡	2.0	0.9	8:20	6:33
9/1‡	1.8	0.5	9:00	6:34
9/2‡	1.1	~0.5	~8:30	6:35

*When [HONO] = 0.5[HONO]_{max}.

†Local sunrise at Riverside ~45 min later because of nearby high ground.

‡Measurements at Claremont.

found that the reaction proceeded relatively rapidly on the surface of their aged Pyrex reaction vessel. The possibility that the observed HONO may be formed directly in the combustion region of sources, should also be considered.

Whatever the source, at sunrise the HONO accumulated throughout the night is rapidly photolysed, and the HONO destruction rates given in Table 1 lead to OH production rates of up to 2×10^7 OH radicals cm⁻³ s⁻¹. By making an estimate of the lifetime of OH soon after sunrise at Riverside

of ~0.1 s (dictated largely by reaction with NO₂), we can conclude that up to $\sim 2 \times 10^6$ OH radicals cm⁻³ could be produced in an approximately hour-long early morning 'pulse' from HONO photolysis, at a time when other photochemical activity is low.

While nitrosamines have not been widely detected in the atmosphere, our direct observations of p.p.b. levels of HONO in the polluted troposphere at night implies that the possibility of their formation¹⁵, at least near sources of precursor amines, should be seriously considered. After sunrise both HONO and any nitrosamines present would be rapidly depleted by photolysis.

Received 26 November 1979; accepted 26 March 1980.

- Pitts, J. N. Jr & Finlayson-Pitts, B. J. *Adv. Envir. Sci. Technol.* **7**, 75-162 (1977).
- Falls, A. H. & Seinfeld, J. H. *Envir. Sci. Technol.* **12**, 1398-1406 (1978).
- Hendry, D. G., Baldwin, A. L., Barker, J. R. & Golden, D. M. *Computer Modeling of Simulated Photochemical Smog* (EPA-600/3-78-059, 1978).
- Carter, W. P. L., Lloyd, A. C., Sprung, J. L. & Pitts, J. N. Jr *Int. J. Chem. Kinet.* **11**, 45-101 (1979).
- Pitts, J. N. Jr *Phil. Trans. R. Soc. A290*, 551-576 (1979).
- Platt, U., Perner, D. & Patz, H. W. *J. geophys. Res.* **84**, 6329-6335 (1979).
- Perner, D. & Platt, U. *Geophys. Res. Lett.* **6**, 917-920 (1979).
- JPL Publication 79-27, *Chemical Kinetic and Photochemical Data for Use in Stratospheric Modelling* (1979).
- Howard, C. J. *J. phys. Chem.* **67**, 5258-5263 (1977).
- Graham, R. A., Winer, A. M. & Pitts, J. N. Jr *Chem. phys. Lett.* **51**, 215-220 (1977).
- Batt, L. *Int. J. Chem. Kinet.* **11**, 977-993 (1979).
- Kaiser, E. W. & Wu, C. H. *J. phys. Chem.* **81**, 1701-1706 (1977).
- McKinnon, I. R., Mathieson, J. G. & Wilson, I. R. *J. phys. Chem.* **83**, 779-780 (1979).
- Stockwell, W. R. & Calvert, J. G. *J. Photochem.* **8**, 193-203 (1978).
- Tuazon, E. C., Winer, A. M., Graham, R. A., Schmid, J. P. & Pitts, J. N. Jr *Envir. Sci. Technol.* **12**, 954-958 (1978).

Correlation of ¹⁸O in precipitation with temperature and altitude

U. Siegenthaler & H. Oeschger

Physics Institute, University of Bern, Bern CH-3012, Switzerland

The ¹⁸O/¹⁶O ratio in precipitation is correlated to climatic parameters, especially temperature. However, the meteorological processes determining the stable isotope composition of precipitation are complex and only partly understood, so that a quantitative interpretation of ancient isotope ratios in precipitation, as recorded in polar ice¹ or in carbonate sediments from lakes², has not yet been possible. For several years we have measured ¹⁸O/¹⁶O ratios on monthly precipitation samples from Switzerland and here summarize these data.

Samples were collected at five stations (circles in Fig. 1) of the network of the Swiss Meteorological Service, so that meteorological data are at hand for correlation with ¹⁸O/¹⁶O results (Table 1). Three of the stations are situated in a valley in the Central Alps (Haslital, Eastern Bernese Oberland): Meiringen at 632 m, Guttannen at 1,055 m and Grimsel at 1,950 m above sea level. Guttannen and Grimsel are respectively 11 and 20 km away from Meiringen. The valley rises from NW to SE and is surrounded by mountains that reach altitudes between 2,000 m and 3,500 m. Bern (572 m) is situated about 70 km WNW of these three stations and away from the direct influence of the mountains. Locarno is on the southern side of the Alps, at an altitude of 366 m, and its climate is markedly different from that of the other stations.

An important factor determining $\delta^{18}\text{O}$ in precipitation is temperature. Dansgaard³ found that on a global scale there is a good correlation for coastal and polar stations between annual means of $\delta^{18}\text{O}$ and of temperature, with an average change in $\delta^{18}\text{O}$ of 0.7‰ per degree Celsius. This global

relation between $\delta^{18}\text{O}$ and temperature can be explained semiquantitatively by considering the formation of precipitation as a Rayleigh process, whereby with continuing formation of precipitation from an air mass the remaining water vapour is progressively being depleted in ¹⁸O because H₂¹⁸O condenses preferably to H₂¹⁶O (refs 3, 4). According to this simple model, $\delta^{18}\text{O}$ in precipitation on the temperature of condensation, but also on the initial conditions in the vapour source region (temperature and isotopic composition). There are further processes influencing $\delta^{18}\text{O}$ which are not taken into account by the Rayleigh model, most significantly admixture of isotopically different water vapour, and ¹⁸O enrichment of the falling raindrops due to evaporation and isotope exchange with ambient vapour^{3,5,6}.

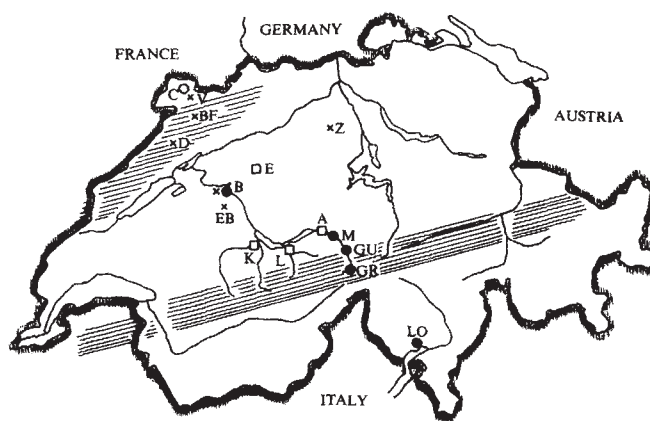


Fig. 1 Map of Switzerland with the locations mentioned in the text. ○, Precipitation sampling stations: B, Bern; M, Meiringen; Gu, Guttannen; Gr, Grimsel; Lo, Locarno; C, Coeuv. ×, springs or groundwater: V, Vendlin; BF, Blanche Fontaine; D, Doux; B, Bern; Eb, Englisberg; Z, Zetzwill. □, Rivers: E, Emme; K, Kander; L, Lütchine; A, Aare. The mountain ranges of the Jura and the Alps are shown schematically as hatched areas.

Table 1 Data for the stations from which precipitation was analysed. The values are based on measurements from the period 1971–78 (1973–78 for Locarno). The temperature coefficient (last line) was obtained as the slope of a regression line of monthly $\delta^{18}\text{O}$ values as a function of temperature

	Grimsel	Guttannen	Meiringen	Bern	Locarno
Altitude above sea level (m)	1,950	1,055	632	572	366
Mean annual temperature ($^{\circ}\text{C}$)	1.2	6.3	7.9	8.8	11.7
Mean annual precipitation (mm per yr)	2,091	1,650	1,271	964	1,945
Weighted mean $\delta^{18}\text{O}$ in precipitation (‰)	−14.4	−12.8	−11.7	−10.0	−8.8
Temperature coefficient for seasonal $\delta^{18}\text{O}$ variations (‰ per degree Celsius)	0.44	0.55	0.53	0.35	0.40

In the following we compare $\delta^{18}\text{O}$ with surface temperature at the site of precipitation sampling. According to the Rayleigh model, condensation temperature should be considered instead of surface temperature, and the initial $\delta^{18}\text{O}$ and temperature of the vapour mass should be taken into account, too. This is, however, not possible because the necessary data are not available.

Long-term mean values of $\delta^{18}\text{O}$ and temperature at different stations (Fig. 2a, large symbols) are relatively well correlated; the corresponding regression line has a slope of 0.56‰ per degree Celsius. $\delta^{18}\text{O}$ is lower at all stations than expected from the globally observed relationship (dashed line), a phenomenon known as the 'continental effect'. Possible reasons for this effect are: (1) the influence of re-evaporated continental water, already depleted in ^{18}O ; (2) (in summer) a larger temperature difference between surface and condensation levels (as supported by higher temperature gradients) on the continent than on the coast.

At Meiringen and Guttannen, $\delta^{18}\text{O}$ is lower than expected from extrapolation from the temperature– $\delta^{18}\text{O}$ relation (or also from the altitude– $\delta^{18}\text{O}$ relation, Fig. 2c) obtained from the other stations. Also, the amounts of precipitation are higher at Guttannen and Meiringen than elsewhere at similar altitudes in this region north of the Alps (compare Meiringen with Bern in Table 1). Both observations can probably be explained by an orographic effect. The two stations are surrounded by high mountains which force the moist air to climb up, so that here the condensation levels are higher and condensation temperatures lower than for other stations at comparable altitudes and mean annual temperature, with the consequence of relatively low $\delta^{18}\text{O}$ and high amounts of precipitation.

In Fig. 2a, weighted mean annual $\delta^{18}\text{O}$ values of the years 1971–78 are compared with arithmetic temperature means (see also Fig. 2b). A correlation cannot be seen for any of our stations. For climatic applications this means that caution is necessary when trying to interpret, for instance, extreme $\delta^{18}\text{O}$ values in annual glacier ice layers in terms of temperature. For this comparison weighted means are taken for $\delta^{18}\text{O}$ because they are adequate for palaeoclimatic isotope records as well as for hydrologic application, while arithmetic means of temperature represent the average climate better than weighted temperature means. If looking for causal relationship, however, weighted annual mean $\delta^{18}\text{O}$ should be compared not with the arithmetic annual temperature mean, but rather with weighted mean temperature during precipitation events. Unfortunately, such weighted temperature means are not readily available. However, plots of $\delta^{18}\text{O}$ against temperature of monthly data (see below) and of daily samples collected during limited time show much scatter, which suggests that the correlation would not significantly improve when using weighted annual temperature means. Furthermore, arithmetic annual means of $\delta^{18}\text{O}$ also show no significant correlation with temperature.

Figure 2b shows the variation of $\delta^{18}\text{O}$ and of temperature (average of all stations) with time. The trends of $\delta^{18}\text{O}$ and of temperature are quite different, which again demonstrates that annual mean $\delta^{18}\text{O}$ is not determined uniquely by temperature. On the other hand, there is quite a good correlation between

different stations (Fig. 2b), with parallel trends for the stations north of the Alps as well as Locarno. From 1971 to 1975, a general decrease of 1–2‰ is observed, then an increase from 1975 to 1977, and again a decrease in 1978. This parallelism indicates that the year-to-year variations of $\delta^{18}\text{O}$ are primarily determined by the large-scale weather situation, rather than by local factors.

In Fig. 2c, the variation of mean $\delta^{18}\text{O}$ values with altitude is plotted (means over whole period of observation). Obviously, this is essentially only a different representation of

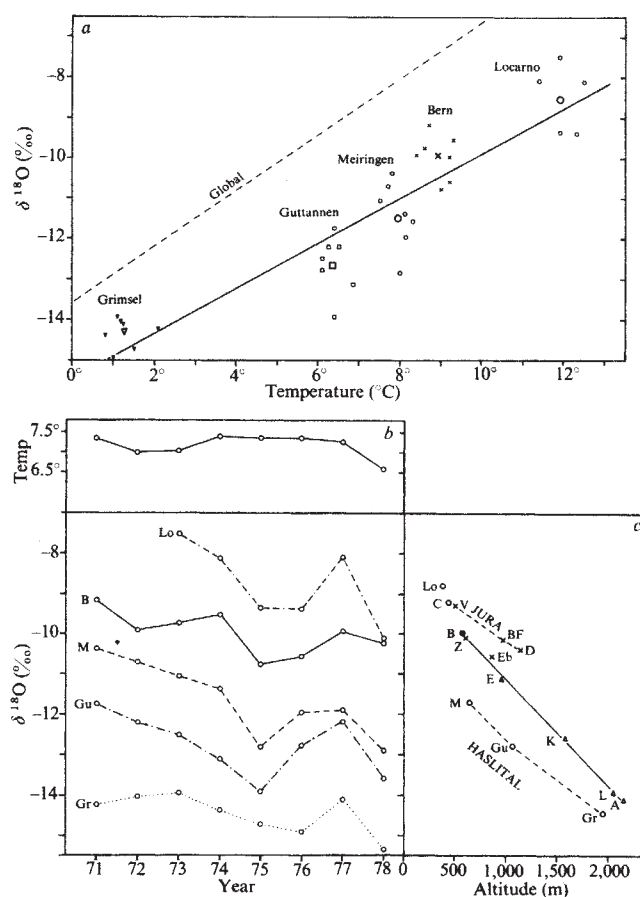


Fig. 2 Mean values of $\delta^{18}\text{O}$, weighted with amount of precipitation. a, $\delta^{18}\text{O}$ values for the years 1971–78 (1973–78 for Locarno) as a function of mean annual air temperatures. Large symbols, mean values for the entire period of observation, with corresponding regression line (solid line), $\delta^{18}\text{O} = 0.56t - 15.5$ ‰. Dashed line, global relationship as obtained by Dansgaard³, $\delta^{18}\text{O} = 0.70t - 13.65$ ‰. b, Time trend of $\delta^{18}\text{O}$ in precipitation for Locarno, Bern, Meiringen, Guttannen and Grimsel. Top, annual mean temperature, average of all five stations. c, $\delta^{18}\text{O}$ as a function of altitude. ○, Precipitation. ×, Springs or groundwater. △, Rivers (see caption to Fig. 1 for names of stations). For rivers and groundwaters, mean altitudes of the catchment basin are given, and $\delta^{18}\text{O}$ data are mean values over at least 18 months.

the $\delta^{18}\text{O}$ -temperature relationship, as temperature and altitude are well correlated in our restricted area. Included are values for rivers and groundwaters from western Switzerland (see map, Fig. 1), because these reflect well the weighted mean $\delta^{18}\text{O}$ of precipitation (see Bern and stations C and V, ref. 7, in Fig. 2c). This is probably also true for the Alpine rivers (Aare, Lütchine, Kander) indicated in Fig. 2c, as evaporation is relatively unimportant at high altitudes, so that most of the annual precipitation must run off as river water.

There is a good correlation between $\delta^{18}\text{O}$ and altitude for locations on the Swiss Plateau (B, Z, Eb, E) and in the western Alps (K, L, A) except those in the Haslital. The corresponding decrease of $\delta^{18}\text{O}$ with altitude (solid line) is -0.26‰ per 100 m. This figure is in the middle of the range of values reported from studies in different regions of Europe⁸⁻¹⁴, which are between -0.16 and -0.4‰ per 100 m (with one exception¹⁵ where -0.5‰ per 100 m was observed). $^{18}\text{O}/^{16}\text{O}$ decreases with altitude also above 2,000 m: mean $\delta^{18}\text{O}$ in precipitation collected at Jungfraujoch (3,572 m) from August 1977 to July 1978 was 1.4‰ lower than at Grimsel, corresponding to a change of -0.09‰ per 100 m between these two stations.

The stations in the Haslital, mainly Meiringen and Guttannen, show rather low isotope values compared with the general trend. This is, of course, the same phenomenon as discussed above in connection with temperature (Fig. 2a), and is presumably due to an orographic uplift of the condensation level, so that the effective altitude of formation of precipitation is higher than in the other regions considered here. Figure 2c shows a more general decrease from west (Jura) to east (Haslital) at constant altitude, comparable on a small scale with a continental effect. The decrease from Jura to Swiss Plateau reflects the rain-shadowing effect of the mountain range of the Jura. Similar phenomena, lower isotopic values on the lee side of a mountain range than on the weather side, were observed at the Sierra Nevada¹⁶ and in Germany⁹. The fact that the elevation (or temperature, respectively) of the site of precipitation sampling is not the only parameter of influence for $\delta^{18}\text{O}$ is important for hydrological applications: stable isotope measurements may be used for estimating the mean altitude of the recharge area of springs^{12,17}, but only if the locally valid $\delta^{18}\text{O}$ -altitude relation is known.

Finally we consider monthly data. Plotted against temperature, the monthly $\delta^{18}\text{O}$ values of a station scatter by about $\pm 3\text{‰}$ around a regression line, which clearly illustrates that temperature alone cannot explain the whole variability of $\delta^{18}\text{O}$. The slopes of the regression lines vary between 0.35 and 0.55‰ per degree Celsius at our stations (last line of Table 1). Figure 3 shows the mean seasonal trends of $\delta^{18}\text{O}$ and temperature for Meiringen, Guttannen and Grimsel. In general, the $\delta^{18}\text{O}$ cycles are about parallel to the seasonal temperature cycles. The difference in $\delta^{18}\text{O}$ between Grimsel and the two stations at lower altitude is, however, clearly higher in summer than in winter, while the temperature difference is relatively constant. The isotopic differences between these stations (in other words, the altitude effect) can be attributed to two main factors. (1) A progressive depletion in H_2^{18}O of the vapour, along with the raining out of a vapour mass during its orographic ascent and cooling. This corresponds to Rayleigh condensation, with progressively higher condensation levels and lower condensation temperatures. (2) ^{18}O enrichment due to evaporation of the falling raindrops. Both processes cause a positive correlation of $\delta^{18}\text{O}$ with temperature (and a negative correlation with altitude). Process (2) depends on the prevailing vapour pressure and is, therefore, more effective at higher temperatures; it is assumed to be negligible for snowflakes because probably only a very thin surface layer can be enriched in ^{18}O (refs 10, 11). Thus the evaporative altitude effect (2) is expected to be larger in summer than in winter. It therefore seems plausible that evaporation of falling raindrops is responsible for the $\delta^{18}\text{O}$ difference between Grimsel and the

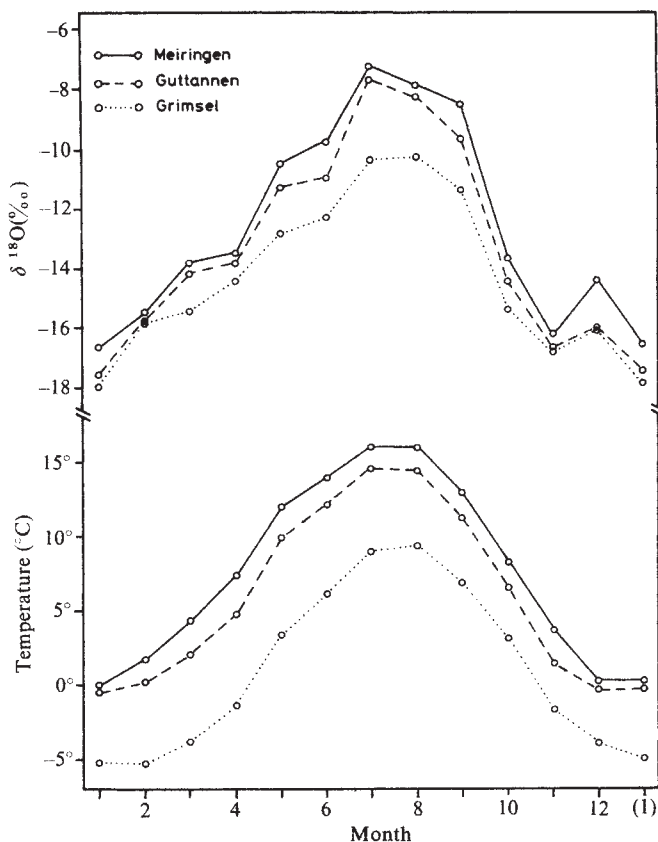


Fig. 3 Monthly $\delta^{18}\text{O}$ values in precipitation and air temperatures, mean values 1971-78. January (1) is shown twice for the sake of continuity.

two lower-altitude stations being larger in summer than in winter, and thus for the different temperature coefficients (Table 1) of these stations.

The temperature- $\delta^{18}\text{O}$ relation is of considerable interest for palaeoclimatic applications. The relation between mean values of different stations (Fig. 2a) partly reflects geographical influences and is, therefore, not directly comparable to climatic temperature and isotopic variations. On the other hand, there is a large range of 'climates' within an annual cycle. It is tempting to speculate that the temperature coefficient found for seasonal variations, $0.45 \pm 0.1\text{‰}$ per degree Celsius, is also valid for long-term temperature changes and therefore for the interpretation of palaeoisotope records in Central Europe. The analogy between seasonal and climatic variations is, however, incomplete, and the temperature coefficient valid for climatic changes may be quite different from that of the seasonal variation. For a better understanding of the processes affecting $\delta^{18}\text{O}$ in precipitation, additional studies covering larger geographical areas will be necessary, considering also the variations of conditions in the source regions of the water vapour.

We thank Mr K. Hänni for analysing the samples, Mr H. Matter for computational aid, and the meteorological observers who collected the samples. This work was supported by the Swiss NSF.

Received 23 July 1979; accepted 21 March 1980.

1. Dansgaard, W., Johnsen, S. J., Møller, J. & Langway, C. C. *Science* **166**, 377-381 (1969).
2. Eichler, U. & Siegenthaler, U. *Boreas* **5**, 109-117 (1976).
3. Dansgaard, W. *Tellus* **16**, 436-468 (1964).
4. Friedman, I., Redfield, A. C., Schoen, B. & Harris, J. *Rev. Geophys. Space Phys.* **2**, 177 (1964).
5. Friedman, I., Machta, L. & Solter, R. *J. geophys. Res.* **67**, 2761-2766 (1962).
6. Ehli, D., Knott, K., Nagel, J. F. & Vogel, J. C. *J. geophys. Res.* **68**, 3775-3780 (1963).
7. Siegenthaler, U., Oeschger, H. & Tongiorgi, E. in *Isotope Hydrology 1970*, 373-385 (IAEA, Vienna, 1970).
8. Dansgaard, W. *Medd. Grønland* **165**, 1-120 (1961).
9. Eichler, R. *Geol. Rdsch.* **55**, 144 (1964).
10. Ambach, W., Dansgaard, W., Eisner, H. & Møller, J. *Tellus* **20**, 595-600 (1968).

11. Dinger, T., Martinec, J., Payne, B. R. & Yen, C. K. in *Isotope Hydrology 1970*, 23–42 (IAEA, Vienna, 1970).
12. Stahl, W., Aust, H. & Douns, A. in *Isotope Techniques in Groundwater Hydrology 1974* Vol. 1, 317–339 (IAEA, Vienna, 1974).
13. Herrmann, A., Rauert, W. & Stichler, W. *Proc. Meet. Sonderforschungsbereich 81*, Technische Universität München, 30–42 (1977).
14. Bortolami, G. C., Ricci, B., Susella, G. F. & Zuppi, G. M. in *Isotope Hydrology 1978* Vol. 1, 327–350 (IAEA, Vienna, 1978).
15. Moser, H. & Stichler, W. in *Isotope Hydrology 1970*, 43–52 (IAEA, Vienna, 1970).
16. Friedman, I. & Smith, G. I. *Science* **169**, 467–470 (1970).
17. Siegenthaler, U. & Schotterer, U. *Gas Wass. Abwass.* **57**, 501–506 (1977).

Accretion and episodic plutonism

Loren A. Raymond & Samuel E. Swanson*

Department of Geology, Appalachian State University, Boone, North Carolina 28608

Episodic calc-alkaline igneous activity in magmatic arcs contrasts with steady-state magmatism accompanying seafloor spreading. Gilluly¹ examined this apparent anomaly in western North America and he and Lanphere and Reed² concluded that igneous activity has been more or less continuous over the entire North American Cordillera since the Triassic. A similar view was held by Pitcher³ for the coastal batholith of Peru. Numerous studies, however, support the view that episodic magmatism characterizes areas of local to regional scale along convergent plate margins^{1,2,4–12}. Grow and Atwater¹³ suggested that increased magmatic activity in Alaska might have been induced by subduction of a spreading ridge. If so, episodic igneous activity could be produced by successive subduction events involving spreading centres⁸. Dickinson⁷ suggested a model involving frictional heating along the Benioff zone, accumulation of magma above that zone, and periodic magmatism triggered by an unknown condition or event. However, he later⁸ proposed that regional episodicity is a function of the interaction between a specific site and a wavering (migrating back and forth) magmatic axis that occasionally passes through the site to produce a magmatic event. Migration of the magmatic axis results from accretion or erosion at the trench, variation in the dip of the subduction zone, and (or) reorganization of plate movements^{8,14–16}. Shaw *et al.*¹⁷ suggest tidal power as a source of mechanical energy which, through periodic variations in shear melting of the mantle, produces magmatic periodicity. In contrast, Gilluly¹ suggests that variations in the properties of the mantle and of the crust of the subducting plate may be responsible for magmatic episodicity. We propose an alternate explanation for local to regional plutonic episodicity. Although this 'episodic accretion plutonism model', is based on data from California, it is applicable to other segments of orogenic belts including southwestern Alaska, southern Japan and western Indonesia. We limit the discussion of igneous activity to plutonism, as geochronological data suggest that volcanism is not simply the surficial expression of plutonism. This possibility is supported by a recent re-evaluation of Andean magmatism¹⁸. We shall discuss the problem of volcanism and its relationship to accretion and plutonism elsewhere.

The episodic accretion–plutonism model requires that plutonic activity within segments of orogenic belts alternates with periods of accretion^{19,20}. We propose that plutonism in magmatic arcs is a function of the subduction of hydrous crustal materials, especially sediments, and that plutonism is terminated by episodic accretion of these hydrous materials in the arc–trench gap adjacent to the magmatic arc. Thus, episodic plutonism is related to episodic accretion.

Episodicity may be discussed in terms of a single segment of an orogenic belt rather than the entire belt. We have chosen the California segment of the Cordilleran Mountain system as the prime example because of an excellent data base. The

study area is limited on the north by the Oregon border, on the south by the Garlock fault, on the west by the San Andreas fault, and on the east by the California–Nevada border (Fig. 1). Selection of the California–Nevada border as a boundary excludes Mesozoic–Cenozoic plutonic rocks exposed to the east²¹. However, because of the limited outcrop area of these rocks, the general lack of detailed data on discordance and resetting of dates, and the similarity in overall duration of magmatic activity between the eastern areas and California, the California data set is representative and is the only one that includes enough closely spaced, unambiguous dates and plutons to be adequately depicted on a small scale map. In the California segment of the Cordilleran Orogenic belt, some volcanic, volcanic–clastic, and sedimentary rocks of the western Sierra Nevada represent an early to middle Mesozoic accretionary wedge. The Franciscan Complex of the Coast Ranges consists of late Mesozoic and early Cenozoic accretionary wedges. The positions of the magmatic arcs of different ages, which all lie within the Sierra Nevada and Klamath Mountains, are shown in Fig. 1.

Documentation of the episodic accretion–plutonism model requires careful dating of both the accreted wedge and the magmatic arc. Figure 2 is a compilation of the data on accretion and plutonic activity for the Mesozoic and earliest Cenozoic eras in the study area. The controversial intrusive

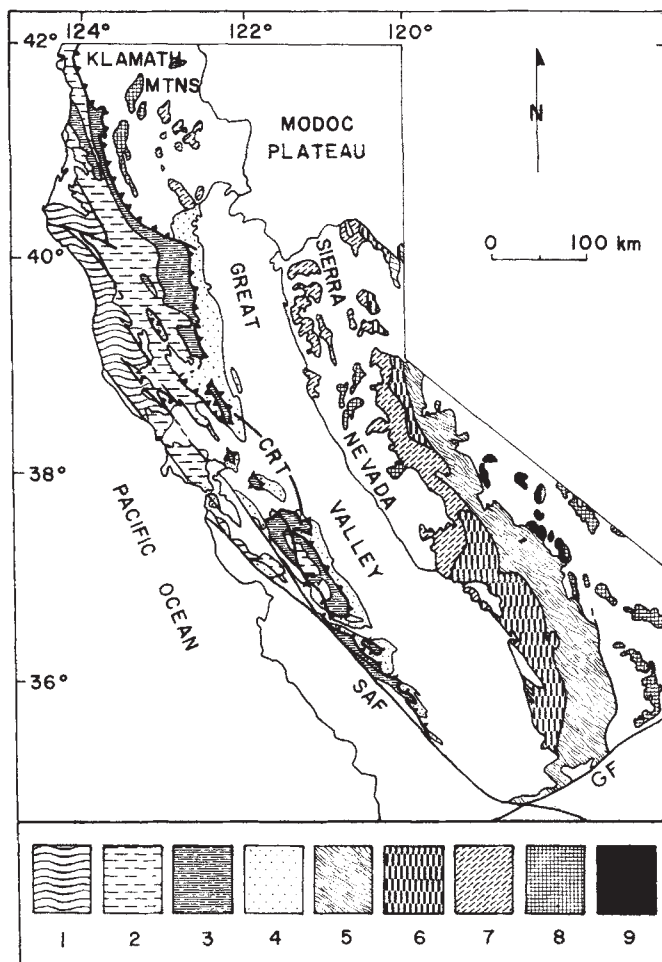


Fig. 1 Map of California showing the distribution of accreted (Franciscan) rocks and granitic rocks intruded during Sierran plutonic episodes. 1, Mid-Cenozoic accretionary wedge; 2, late Cretaceous accretionary wedge; 3, early Cretaceous accretionary wedge; 4, Great Valley sequence (arc-trench gap deposits); 5, late Cretaceous plutons; 6, early Cretaceous plutons; 7, late Jurassic plutons; 8, early Jurassic plutons; 9, late Triassic plutons. SAF, San Andreas Fault; GF, Garlock Fault; CRT, Coast Range Thrust (modified from refs 6, 25, 30).

*Present address: Division of Geoscience, Geology/Geophysics Program, University of Alaska, Fairbanks, Alaska 99701.

epochs of Evernden and Kistler^{5,6,10} are used as a basis for the age of maximum pluton emplacement in the arc.

Dating episodes of accretion is more difficult than dating plutons. Mapping of accreted terrains reveals that they are composed of structurally and stratigraphically complex, generally unfossiliferous, rock assemblages. Pelagic sediments deposited on a plate moving towards a subduction zone may be succeeded by arc/continent derived clastic sediments deposited on submarine fans and by submarine landslide deposits (olistostromes) and proximal fan facies deposited in the trench. Trench slope basin deposits may cap the stratigraphic sequence before or during deformation and tectonic mélangé formation in the subduction zone. As sedimentation may be continuous over the period of subduction activity, fossil dates on the entire accreted wedge will span the convergence interval. Yet, mapping and palaeontological studies reveal that accreted terrains may be subdivided into time restricted stratigraphic/structural units. Because sedimentation is terminated by the onset of accretion and/or subduction, the upper age limit given by the fossils in any stratigraphic/structural unit is considered to be the approximate age of accretion. In the Franciscan Complex, numerous studies²²⁻²⁹ on fossils contained in major structural units^{22,23,27,30} allow estimates of the age of accretion of these units (Fig. 2). The ages of major accreted structural units in the Sierra Nevada are less well known, but one Mesozoic mass

may be exposed³¹. Some accreted rocks may also be concealed beneath the Great Valley. To the north, in the Klamath Mountains, accreted rocks of Triassic-Jurassic age are present in the so-called 'western Palaeozoic and Triassic belt'³² (Fig. 2).

Metamorphic dates on rocks of the accreted wedge, especially blueschist facies rocks, also mark times of accretion. That blueschist dates mark accretion events may seem to contradict the common view that blueschists are formed continuously during subduction³³. Although blueschist facies rocks are probably formed continuously, it is only when rocks subducted to depths of <30 km are uplifted rapidly, for example during upward rotation of accreted wedges³⁴, that these rocks are preserved. Blueschist K/Ar dates mark the time at which accreted wedges pass upwards through the argon retention isograd. Thus, blueschist dates generally fall at the end of accretion events (Fig. 2), reflecting the time required for uplift and upward rotation of major structural/stratigraphic units.

Peak dates of metamorphism used in Fig. 2 (columns 2 and 4) are based on new dates obtained by D. Krummenacher and D. L. Martin for L. A. Raymond, plus additional dates from the literature. The meaning of metamorphic dates on accreted rocks in the Sierran foothills is equivocal, but we have used the few dates available in Behrman³⁵. Dates on blueschist facies Coast Range rocks from refs 33, 36-41 were considered in constructing Fig. 2, but special attention was given to the re-evaluation of dates by Lanphere and others⁴¹.

Coincidence of Coast Range blueschist dates and dates of termination of sedimentation in major tectono-stratigraphic units (accretion dates) with gaps in the plutonic history of the Sierra Nevada shown in Fig. 2 provides the basis for the episodic accretion-plutonism model. (Because metamorphic dates are dates of cooling, it is surprising that the coincidence is as marked as it appears in Fig. 2. The problems of dating low grade rocks and the variations expected are discussed elsewhere^{40,41}. It is expected, for example, that metamorphic dates will tend to overlap the beginning of the next plutonic episode, as is the case with the 120 Myr date in the Franciscan rocks.) In addition, metamorphic dates on accreted Sierran and Klamath rocks fall in the gaps between episodes of plutonism. Together these data indicate that times of accretion in the arc-trench gap are not times of plutonism in the magmatic arc. This suggests that plutonism is controlled by subduction (and subsequent dehydration) of sediments.

Episodic accretion and plutonism are not restricted to northern California, although the timing of events in other segments of orogenic belts is generally less well known. In southwestern Alaska, accretion of a blueschist facies and lower grade terrains at ~190, 110, 50 and 20 Myr is interspersed with plutonic events dated at 195-185, 176-145, 83-58 and 28-26 Myr⁴²⁻⁴⁸. In western Indonesia, mid-Cretaceous and Miocene intrusive events were followed by late Cretaceous and Miocene accretion events⁴⁹. Accretion events in south-west Japan during the early Cretaceous (?) and Oligocene are followed by plutonic events of mid-Cretaceous to early Tertiary age (130-60 Myr) and late Miocene to Recent age (16-0 Myr)⁵⁰.

Clearly, careful dating of events along convergent margins will be required to test the episodic accretion-plutonism model. However, available data are adequate to justify use of this model in evaluating the tectonic history of plate margins preserved in the continental geologic record.

We thank R. Kistler and G. Plafker for discussions, R. W. Kistler and J. D'Allura for reviews, and M. Raymond, C. Muirhead, and J. Webb for assistance in preparing the manuscript. Research was supported by the Division of Earth Sciences, NSF grant EAR 76-06062 (to L. A. Raymond), by Appalachian State University and by Southern Oregon State College.

Received 7 February; accepted 1 April 1980.

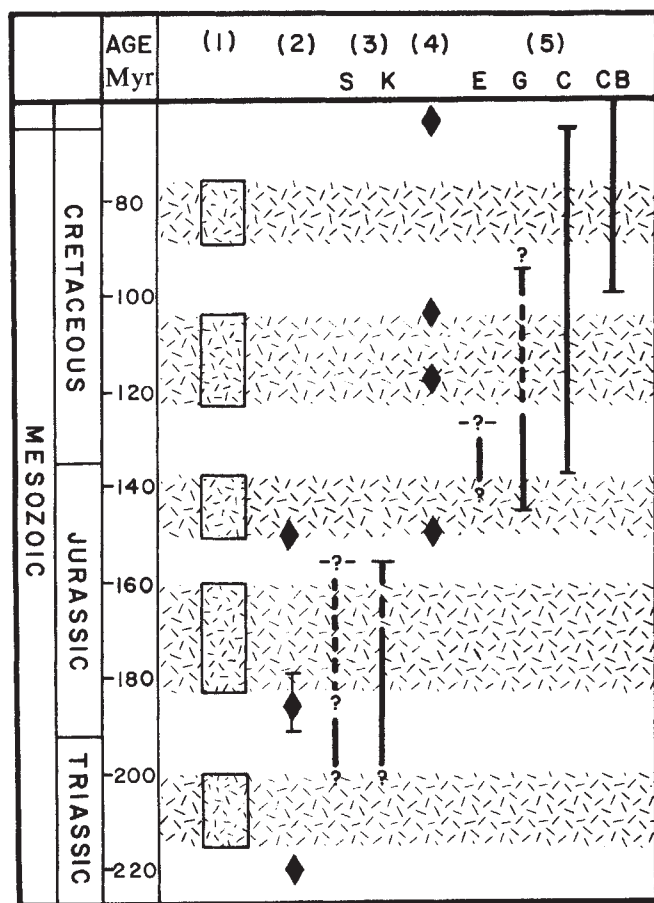


Fig. 2 Correlation chart showing: (1) Sierran intrusive episodes (after refs 5, 6, 10); (2) Sierran and Klamath metamorphic events (diamonds) (after refs 35, 41, 51); (3) Sierran (S) and Klamath (K) sedimentation episodes (after refs 31, 32); (4) peak dates of blueschist facies metamorphism in the Coast Range (Franciscan) after refs 36-41 and L.A.R., unpublished data; (5) Franciscan sedimentation episodes: E, Eastern Belt; C, Central Belt; G, the Geysers block; CB, Coastal Belt (after refs 22, 23, 25, 26, 52-55).

Time scales from ref. 56.

1. Gilluly, J. *Bull. geol. Soc. Am.* **84**, 499–514 (1973).
2. Lanphere, M. S. & Reed, B. L. *Bull. geol. Soc. Am.* **84**, 3773–3782 (1972).
3. Pitcher, W. S. *Bull. geol. Soc. Lond.* **131**, 587–591 (1975).
4. Kawano, Y. & Ueda, Y. *Tectonophysics* **4**, 523–530 (1967).
5. Evernden, J. F. & Kistler, R. W. *Prof. Pap. U.S. geol. Survey* **623** (1970).
6. Kistler, R. W., Evernden, J. F. & Shaw, H. R. *Bull. geol. Soc. Am.* **82**, 853–868 (1971).
7. Dickinson, W. R. *Rev. geophys. Space Phys.* **8**, 813–860 (1970).
8. Dickinson, W. R. *Am. J. Sci.* **272**, 551–576 (1972).
9. Crowder, D. F., McKee, E. H., Ross, D. C. & Krauskopf, K. B. *Bull. geol. Soc. Am.* **84**, 285–296 (1973).
10. Kistler, R. W. *A Rev. Earth planet. Sci.* **2**, 403–418 (1974). Kistler, R. W. in *Pacific Coast Paleogeography Symp.* Vol. 2, 75–84 (Soc. Econ. Paleontol. Miner., Pacific Section, 1978).
11. Bussell, M. A., Pitcher, W. S. & Wilson, P. *Can. J. Earth Sci.* **13**, 1020–1030 (1976).
12. Armstrong, R. L., Taubeneck, W. H. & Hales, P. O. *Bull. geol. Soc. Am.* **88**, 397–411 (1977).
13. Grow, J. A. & Atwater, T. *Bull. geol. Soc. Am.* **81**, 3715–3722 (1970).
14. Coney, P. J. *Am. J. Sci.* **272**, 603–628 (1972).
15. Coney, P. J. & Reynolds, S. J. *Nature* **270**, 403–406 (1977).
16. Keith, S. B. *Geology* **6**, 516–521 (1978).
17. Shaw, H. R., Kistler, R. W. & Evernden, J. F. *Bull. geol. Soc. Am.* **82**, 869–896 (1971).
18. Thorpe, R. S. & Francis, P. W. in *Origin of Granite Batholiths*, 65–75 (Shiva, Orpington, 1979).
19. Coleman, R. G. *24th int. geol. Congr.* Sect. 2, 19–26 (1972).
20. Prince, R. A. & Kulm, L. D. *Bull. geol. Soc. Am.* **86**, 1639–1653 (1975).
21. Cross, T. A. & Pilger, R. H. Jr *Am. J. Sci.* **278**, 865–902 (1978).
22. Bailey, F. H., Irwin, W. P. & Jones, D. L. *Bull. Calif. Div. Mines Geol.* **183** (1964).
23. Blake, M. C. Jr & Jones, D. L. *Spec. Pub. Soc. Econ. Paleontol. Miner.* **19**, 345–357 (1974).
24. Pessagno, E. A. Jr *Geology* **1**, 153–156 (1973).
25. McLaughlin, R. J. & Pessagno, E. A. Jr *J. Res. U.S. geol. Survey* **6**, 715–726 (1978).
26. Evitt, W. R. & Pierce, S. T. *Geology* **3**, 433–436 (1975).
27. Berkland, J. O., Raymond, L. A., Kramer, J. C., Moores, E. M. & O'Day, M. *Bull. Am. Ass. petrol. Geol.* **56**, 2295–2302 (1972).
28. O'Day, M. & Kramer, J. C. in *Geol. Soc. Sacramento Guidebook* **27**, 51–56 (1972).
29. Maxwell, J. C. *Bull. geol. Soc. Am.* **85**, 1195–1204 (1974).
30. Jones, D. L., Blake, M. C. Jr, Bailey, E. H. & McLaughlin, R. J. *Tectonophysics* **47**, 207–222 (1978).
31. Schweickert, R. A. in *Pacific Coast Paleogeography Symp.* Vol. 2, 361–384 (Soc. Econ. Paleontol. Miner. Pacific Section, 1978).
32. Irwin, W. P., Jones, D. L. & Pessagno, E. A. Jr *Geology* **5**, 557–562 (1977).
33. Suppe, J. *24th Int. geol. Congr.* Sect. 3, 552–559 (1972).
34. Karig, D. E. & Sharman, G. F. *Bull. geol. Soc. Am.* **86**, 377–389 (1975).
35. Behrman, P. S. in *Pacific Coast Paleogeography Symp.* Vol. 2, 337–348 (Soc. Econ. Paleontol. Miner. Pacific Section, 1978).
36. Suppe, J. *Bull. geol. Soc. Am.* **80**, 135–142 (1969).
37. Coleman, R. G. & Lanphere, M. A. *Bull. geol. Soc. Am.* **82**, 2397–2412 (1971).
38. Suppe, J. & Armstrong, R. L. *Am. J. Sci.* **272**, 217–233 (1972).
39. Lehman, D. H., Gucwa, P. R., Fritz, D. & McDowell, F. W. *Abstr. Prog. geol. Soc. Am.* **9**, 452 (1977).
40. Suppe, J. & Foland, K. A. in *Pacific Coast Paleogeography Symp.* Vol. 2, 431–452 (Soc. Econ. Paleontol. Miner., Pacific Section, 1978).
41. Lanphere, M. A., Blake, M. C. Jr & Irwin, W. P. *Am. J. Sci.* **278**, 798–815 (1978).
42. Forbes, R. B. & Lanphere, M. A. *J. geophys. Res.* **78**, 1383–1386 (1973).
43. Reed, B. L. & Lanphere, M. A. *Bull. geol. Soc. Am.* **84**, 2583–2610 (1973).
44. Moore, J. C. in *The Geology of Continental Margins*, 811–816 (Springer, Berlin, 1974).
45. Carden, J. R., Connelly, W., Forbes, R. B. & Turner, D. L. *Geology* **5**, 529–533 (1977).
46. Plafker, G., Jones, D. L. & Pessagno, E. A. Jr. in *Circular U.S. geol. Survey* **751-B**, 41–42 (1977).
47. Plafker, G. et al. *Open-file Rept. U.S. geol. Survey* 78–490 (1978).
48. Hudson, T. *Geology* **7**, 230–234 (1979).
49. Katili, J. A. in *The Western Pacific: Island Arcs, Marginal Seas, Geochemistry*, 287–306 (Crane, Russak, Canberra, 1973).
50. Matsumoto, T. & Kimura, T. in *Mesozoic-Cenozoic Orogenic Belts*, 513–542 (Scottish Academic Press, Edinburgh, 1974).
51. Hotz, P. E., Lanphere, M. A. & Swanson, D. A. *Geology* **5**, 659–663 (1977).
52. Wachs, D. & Hein, J. R. *Geology* **3**, 29–33 (1975).
53. Berkland, J. O. *24th Int. geol. Congr.* Sect. 3, 99–105 (1972).
54. Gucwa, P. R. *Geology* **3**, 105–108 (1975).
55. Bachman, S. B. in *Pacific Coast Paleogeography Symp.* Vol. 2, 419–430 (Soc. Econ. Paleontol. Miner. Pacific Section 1978).
56. van Hinte, J. E. *Bull. Am. Ass. petrol. Geol.* **60**, 269–287 (1976); 489–497 (1976).

Caledonian Sm–Nd ages and a crustal origin for Norwegian eclogites

W. L. Griffin

Mineralogisk-Geologisk Museum, Oslo, Norway

H. K. Brueckner

Queens College of the City University of New York, Flushing, New York 11367 and Lamont-Doherty Geological Observatory, Palisades, New York 10964

Sm–Nd dating of minerals in several Norwegian country-rock eclogites reported here indicates that the rocks crystallized ~425 Myr ago. The isotopic composition of both Nd and Sr in these eclogites suggests that they formed from much older crustal rocks. The high-pressure metamorphism of which the eclogites are relics probably involved subduction of the continental Baltic plate as it was overridden by the Greenland plate during the Caledonian orogeny.

Eclogites of the Caledonian orogeny in western Norway are usually divided¹ into two types: mafic and ultramafic layers or lenses within garnet-free ultramafic bodies ('internal' or 'Type A' eclogites); and mafic boudins, layers and larger bodies enclosed directly within Precambrian gneisses ('external', 'country-rock', or 'Type B' eclogites). The origin(s) and time(s) of formation of the latter eclogites are controversial. Several workers regard them as products of progressive regional metamorphism in essentially their present stratigraphic and tectonic position. Evidence for this view¹ includes: (1) zoning in minerals, indicating growth with increasing temperature (*T*) and pressure (*P*); (2) preservation of lower-grade minerals in the cores of garnet grains; (3) field and geochemical indications that the protoliths had low-*P* origins; and (4) a regional gradient in maximum metamorphic *P* and *T* with highest values near the coastline. Krogh² proposed that the eclogites formed by metamorphism of a variety of protoliths during subduction of the continental crust. On the basis of available radiometric data^{3–5} he suggested that this event was part of the Svecofennian (1,600–1,800 Myr) orogeny.

An alternative view holds that the eclogites and ultramafic rocks formed in the mantle and were later introduced into the crust as solid fragments^{6,7}. Most proponents of this model favour a Caledonian origin for eclogite formation.

Neither model is necessarily time-dependent; nor does the *in situ* model preclude the incorporation of mantle fragments into the crust during subduction. However, the orogenic processes implied by the two schemes are quite different. Knowledge of the timing and mechanism of eclogite formation in this region is especially important for our understanding of the Caledonian orogeny.

We have attempted to solve these problems by Sm–Nd dating of minerals from several country-rock eclogites. The garnet-clinopyroxene pair is ideal for Sm–Nd dating, as the two phases fractionate Sm and Nd more efficiently than most rock-forming minerals. Furthermore, as Sm and Nd are nearly identical in their crystal-chemical behaviour, the isotopic system should be resistant to metamorphic effects short of total recrystallization of the host minerals. This argument, however, has not yet been rigorously tested.

Analyses were done at Lamont-Doherty Geological Observatory using methods already described⁸. The isotopic data are normalized to ¹⁴⁶Nd/¹⁴²Nd = 0.7219 and ⁸⁶Sr/⁸⁸Sr = 0.1194. The ages obtained are relatively precise, because of the large difference in the Sm/Nd ratio between garnet and clinopyroxene (Table 1), although some individual analyses are less precise than desired. The uncertainties reflect difficulties with the separation of Nd, especially from garnets. The results in Table 1 show three important features.

(1) Five out of six garnet-clinopyroxene pairs give two-point isochron ages between 407 and 447 Myr. One mineral pair is significantly older (887 Myr).

(2) A wide range in initial ¹⁴³Nd/¹⁴⁴Nd ratios (0.51162–0.51244) suggests that eclogites formed from various protoliths. If the protoliths of samples K-6, N-16 and 1428 separated from a mantle of chondritic uniform reservoir (chur) type⁹, they did so long before the Caledonian orogeny. Samples 5/79A and S19 are ambiguous in this respect, as their time-integrated Sm/Nd ratios are nearly chondritic.

(3) *T*_{chur} model ages⁹ for most individual minerals are discordant and geologically meaningless, as would be expected if metamorphism took place long after separation of the protolith from the mantle.

The ages reported here are similar to U–Pb ages on zircon from the Hareidland eclogite (400–418 Myr; compare with sample N-16)¹⁰, to Rb–Sr whole-rock ages on late/post-tectonic granitic dykes near Kristiansund (380 Myr; compare

Table 1 Sm–Nd analytical data from Norwegian eclogites (ages in Myr)

Sample	Sm Phase (p.p.m.)	Nd (p.p.m.)	$^{147}\text{Sm}/^{144}\text{Nd}$	$^{143}\text{Nd}/^{144}\text{Nd}$	T_{chur}	Cpx–gnt age (Myr)	$^{143}\text{Nd}/^{144}\text{Nd}_i$	$^{87}\text{Sr}/^{86}\text{Sr}_i^*$
K-6	cpx 10.49 ± 1	2.890 ± 5	0.1656	0.512065 ± 46	3081			
	gnt 1.120 ± 2	1.397 ± 4	0.7499	0.513621 ± 48	269	407 ± 24	0.511624 ± 73	
5/79	cpx 2.096 ± 5	1.126 ± 2	0.3231	0.512980 ± 40	401			
	gnt 0.7507 ± 8	1.906 ± 5	1.528	0.516316 ± 50	420	423 ± 12	0.512086 ± 65	0.71646 ± 40
U-19	cpx 4.099 ± 12	1.583 ± 3	0.2306	0.512418 ± 64	915			
	gnt 0.4717 ± 14	0.4034 ± 16	0.5146	0.514051 ± 46	671	877 ± 59	0.511092 ± 153	0.71756 ± 8†
S-19	cpx 10.13 ± 14	2.981 ± 5	0.1769	0.512560 ± 36	730			
	gnt 0.4874 ± 26	0.8423 ± 80	1.040	0.515087 ± 72	429	447 ± 20	0.512056 ± 44	0.71496 ± 14
1428	cpx 5.170 ± 16	2.652 ± 10	0.3085	0.513282 ± 34	852			
	gnt 1.570 ± 2	1.507 ± 4	0.6081	0.514102 ± 40	538	418 ± 11	0.512437 ± 110	
N-16	cpx 3.536 ± 3	1.105 ± 7	0.1879	0.512887 ± 36	<0			
	gnt 0.5485 ± 16	1.308 ± 4	1.436	0.516340 ± 150	455	423 ± 30	0.512367 ± 32	0.70589 ± 4†

K-6, Pod in gneiss, Frei, south-east of Kristiansund. Cpx + gnt + opx + biot + qtz.

5/79, Small lens in gneiss, Raudberg, north of Måløy. Cpx + gnt + qtz.

U-19, Scree sample, central Otrøy. Cpx + opx + gnt + bio + qtz. Gnt exsolution in opx.

S-19, Small pod of pegmatitic eclogite, Selje Church, north-east of Måløy. Cpx + gnt + opx.

1428, Thick layer adjoining extensive marble horizon, Tverfjell, north of Molde. Cpx + gnt + amph + qtz.

N-16, Very large body, Hareidlandet. Cpx + gnt + qtz; cpx largely symplectitized.

*Recalculated to 424 Myr BP. †Rb and Sr concentrations not measured, $^{87}\text{Rb}/^{86}\text{Sr}$ assumed equal to 0.0012.

with sample K-6)⁴, and to Rb–Sr and K–Ar mineral ages from the gneiss region¹¹. The simplest interpretation is that these ages reflect isotopic homogenization between minerals as a result of major recrystallization of the eclogite protoliths, and presumably of the enclosing Precambrian country rocks, during the Caledonian metamorphism. The Sm–Nd ages are up to 70 Myr older than most Rb–Sr and K–Ar mineral dates (380–410 Myr). This suggests that the Sm–Nd system closed sooner than the Rb–Sr and K–Ar systems, and that these ages more closely approximate the peak of Caledonian metamorphism.

It might be argued that the Caledonian Sm–Nd ages reflect a resetting of Sm–Nd systems in previously formed eclogites. This would require that the Sm–Nd systems are as easily reset as, for example, K–Ar systems, which seems unlikely on present evidence. Otherwise, this metamorphism must have taken place in conditions in which the eclogite minerals remained stable; the conclusion must still be that eclogite-facies metamorphism occurred during the Caledonian orogeny. This event, rather than a Precambrian one, thus seems to be largely responsible for the mineral assemblages of the basal gneisses of western Norway.

Clinopyroxenes from country-rock eclogites have $^{87}\text{Sr}/^{86}\text{Sr}$ ratios at 424 Myr that range from 0.7039 to 0.7176 (ref. 12) (see also Table 1). These ratios suggest that the eclogite protoliths, which have low Rb/Sr ratios (<0.1), had a premetamorphic crustal history during which they acquired radiogenic Sr, and Ar (ref. 5), from the enclosing gneisses. This is true even of samples (5/79A, S19) for which Sm/Nd data are ambiguous. A T_{chur} model age of ~1,600 Myr can be calculated for N-16, using available NAA analyses (Sm/Nd = 0.39) and the ($^{143}\text{Nd}/^{144}\text{Nd}$)₀ value in Table 1. The eclogite protoliths may thus have been exposed to the effects of both the Svecofennian and Sveconorwegian (~1,000 Myr) orogenies while they were in the crust.

The anomalously old age (877 Myr) of sample U-19 may reflect the effects of these earlier metamorphic events. Some eclogite garnets contain inclusions of garnet–hornblende–biotite plagioclase–quartz assemblages that formed in almandine–amphibolite facies conditions. The cores of many eclogite garnets may be relicts of garnet–amphibolite assemblages formed during a Precambrian metamorphic event. The 877 Myr ‘age’ of sample U-19 could therefore reflect the Precambrian age of the garnet core and the Caledonian age of the garnet rim. Studies of eclogites with strongly zoned garnets should clarify this point.

The combined Nd and Sr isotopic data thus suggest that the protoliths of the eclogites studied existed as crustal rocks long before they were metamorphosed to eclogites during the Caledonian orogeny. There is no difference, in this respect, between eclogites with and without orthopyroxene or between those with fresh and symplectitized omphacite. These data are inconsistent with introduction of the eclogites as solid bodies of mantle material, during the Caledonian orogeny. They do support Krogh’s² model of *in situ* metamorphism of various protoliths during subduction of a slab of continental crust. However, the Sm–Nd ages suggest that the actual subduction occurred in Caledonian time.

The samples studied here have all been subjected to temperatures of 700–800 °C and pressures of 14–18 k bar (ref. 2). We suggest that these conditions reflect depression of the Baltic plate as it was overridden by the Greenland (North American) plate along a Himalayan-type suture during the Caledonian orogeny. The Sm–Nd ages probably record a blocking temperature close to the peak of metamorphism. Breakdown assemblages in the eclogites reflect a nearly adiabatic decompression after the metamorphic peak^{2,13}. The sequence of subduction and rapid uplift is similar to that proposed for the eclogite-bearing terrain of the western Alps¹⁴. The Alpine uplift was accompanied by widespread retrograde metamorphism (prasinization); similar retrogression may have produced the present mineralogy of much of the gneiss enclosing the eclogites.

Sm–Nd dating can give very precise ages for the formation of many eclogite assemblages. Further work on these high-pressure relics will provide detailed data on the processes of subduction and obduction in the Caledonian mountain belt and in other orogenic zones that contain eclogites.

We thank R. K. O’Nions for facilities, P. J. Hamilton for technical instruction, and M. Leblang and S. Goldstein for assistance. D. A. Carswell supplied several samples. This research is supported by the Norwegian Council for Science and Humanities and the NSF grant EAR 73-00601.

Received 6 February; accepted 20 March 1980.

1. Bryhni, I., Krogh, E. & Griffin, W. L. *N. Jb. Miner. Abh.* **130**, 49–68 (1977).
2. Krogh, E. *Nature* **267**, 17–19 (1977).
3. Brueckner, H. K. *Norsk Geol. Tidsskr.* **59**, 141–153 (1979).
4. Pidgeon, R. T. & Råheim, A. *Norsk Geol. Tidsskr.* **52**, 241–256 (1972).
5. McDougall, I. & Green, D. H. *Norsk Geol. Tidsskr.* **44**, 183–196 (1964).
6. O’Hara, M. J., Richardson, S. W. & Wilson, G. *Contr. Miner. Petrol.* **32**, 48–68 (1971).
7. Lappin, M. A. & Smith, D. C. *J. Petrol.* **19**, 530–584 (1978).
8. O’Nions, R. K., Hamilton, P. J. & Evensen, N. M. *Earth planet. Sci. Lett.* **34**, 13–22 (1977).

9. DePaulo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 249–252 (1976); **3**, 743–746 (1976).
10. Krogh, T. E., Mysen, B. O. & Davis, G. L. *Carnegie Inst. Wash. Yearb.* **73**, 575–576 (1975).
11. Brueckner, H. K. *Am. J. Sci.* **267**, 1195–1212 (1969).
12. Brueckner, H. K. *Contr. Miner. Petrol.* **60**, 1–15 (1977).
13. Carswell, D. A., Krogh, E. & Griffin, W. L. (in preparation).
14. Ernst, W. G. *Fortschr. Miner.* **54**, 192–222 (1977).

Mantle composition derived from the composition of lherzolites

S. Maaløe & R. Steel

Department of Geology, Realfagbygget, University of Bergen, 5014 Bergen, Norway

Bickle *et al.*¹ recently suggested that the major element composition and rare earth element (REE) content of komatiites indicate a lherzolitic composition for the mantle, and that the REE mantle/chondrite ratio must have a value of ~ 1.0 . The $^{143}\text{Nd}/^{144}\text{Nd}$ ratios for basalts show that the mantle source for basalts has a REE distribution similar to that of chondritic meteorites². The REE lherzolite/chondrite ratios display some variation, but the ratios obtained indicate an average value of about 1.0 (ref. 3), and the REE content of lherzolite is thus consistent with lherzolitic composition for the mantle. The major element composition of the mantle is very important for both geophysics and igneous petrology and we consider here the relevance of the major element composition of lherzolites to the composition of the mantle.

A recent compilation of 384 major element analyses of spinel lherzolite nodules from continental and oceanic regions defines trends for both nodules which are approximately linear and coincident⁴. Why the analyses should form such a well defined trend is unclear as the composition of a module could be controlled by various processes. However, as the nodules represent fragments of the uppermost mantle⁵, the lherzolitic material of the nodules may have resulted from an ascending convection current or plume which generated basalt. If this assumption is correct then the trend of the lherzolites should be related to the process of primary magma generation. This possibility can be checked by comparing the composition of the primary magma and the trend. This comparison will only

Table 1 Major element compositions

	1	2	3	4	5	6
SiO ₂	47.76	50.52	49.50	60.4	44.20	44.58
Al ₂ O ₃	12.06	10.29	16.53	15.7	2.05	2.43
FeO	8.96	9.07	8.46	7.2	8.29	8.27
MgO	17.78	17.78	8.50	3.9	42.21	41.18
CaO	11.20	10.65	12.48	5.8	1.92	2.08
Na ₂ O	1.31	0.70	2.13	3.2	0.27	0.34
K ₂ O	0.03	0.00	0.11	2.5	0.06	0.11
MnO	0.12	0.17	0.08	0.1	0.13	0.13
TiO ₂	0.59	0.36	1.03	1.0	0.13	0.15
P ₂ O ₅	0.05	—	0.06	0.2	0.03	0.03
Cr ₂ O ₃	—	—	—	—	0.42	0.41
NiO	—	—	—	—	0.26	0.25
Sum	99.81	99.54	98.88	100.0	99.95	99.96

1. Primary abyssal tholeiite¹¹.
2. Extrapolated composition for lherzolite trend at 17.78% MgO (ref. 4).
3. Abyssal tholeiite with 8.5% MgO (ref. 8).
4. Average for continental crust¹³.
5. Average of lherzolite, upper mantle composition⁴.
6. Total average for crusts and lherzolite calculated in this paper, model composition for the primitive mantle.

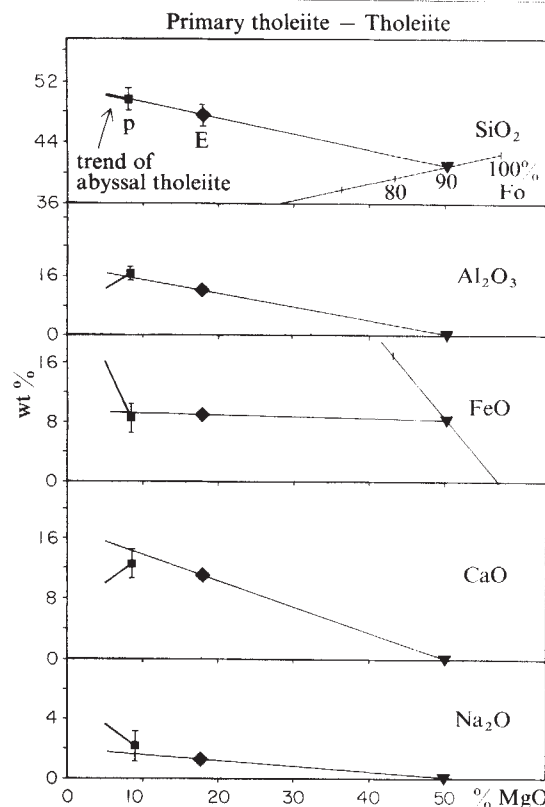


Fig. 1 The olivine control line for primary abyssal tholeiite. The trend for abyssal tholeiite controlled by plagioclase and pyroxene fractionation is shown by a heavy line⁸. The composition 'p' is the most magnesian-rich composition on this trend. Within experimental error the control line passes through p, suggesting that the composition p may be derived from the primary abyssal tholeiite by olivine fractionation.

be possible if the composition of the primary magma generated from the mantle can be reasonably well estimated. We assume here that the primary composition should be that of primary abyssal tholeiite, because these are generated in far larger volumes than any other type of basalt⁶.

The composition of primary abyssal tholeiite has been estimated in various ways. Investigators of abyssal tholeiite have suggested that the most magnesian-rich tholeiites are primary, the results being within the range 9–20% MgO (refs 7–9). This approach does not necessarily give the correct result, and the estimates obtained from ophiolite complexes must be considered more representative. Malpas¹⁰ and Elthon¹¹ have calculated the average composition of the primary abyssal magma by this method and their results are in excellent agreement, the MgO content in both cases being 18% MgO.

The composition of the primary magma estimated by Elthon, is shown in Table 1. The various ophiolite complexes have a similar structure although each displays individual features¹². The general relevance of a primary composition based on a study of a single complex is therefore questionable. However, it is possible to check the composition proposed by Elthon¹¹. The trend of abyssal tholeiites is divided into two parts, a magnesian-rich section controlled by olivine fractionation, and a low magnesian section from 9 to 4% MgO, controlled mainly by plagioclase fractionation⁸. Olivine is the dominant cumulus phase in the ultramafic part of the ophiolite complexes. The fractionation of olivine from the primary magma should therefore yield abyssal tholeiite with $\sim 9\%$ MgO, provided its composition is correct. Figure 1 shows that the estimated composition may be considered

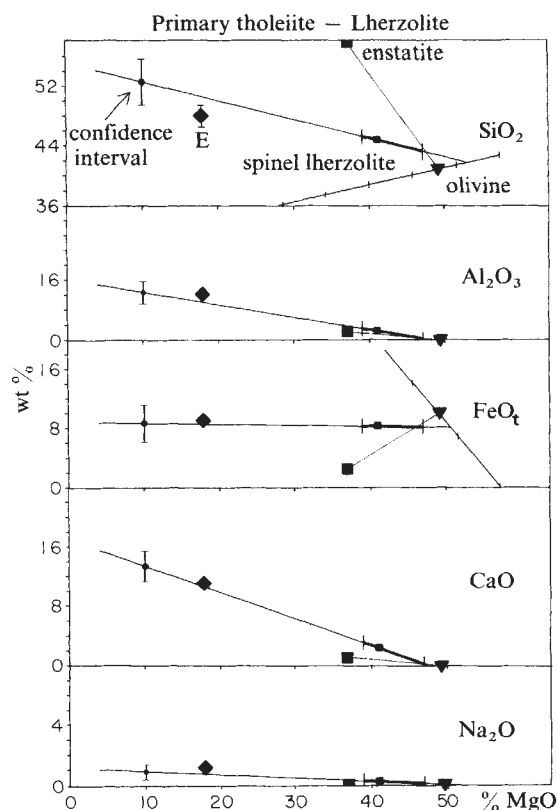


Fig. 2 The compositional relationship between the trend for spinel lherzolites and the primary composition for abyssal tholeiite proposed by Elthon¹¹. There is overlap between the confidence intervals of the trend and the primary composition, the overlap is smallest for SiO_2 . This suggests that the trend of spinel lherzolite is controlled by the generation of primary abyssal tholeiite. The olivine and enstatite compositions are in equilibrium near the solidus of a lherzolite at 20 kbar. The line joining these two compositions will pass through the composition of the residuum formed at this pressure. Accepting that a residuum consists of enstatite and olivine the primary abyssal tholeiite has been formed by $\sim 20\%$ partial melting of a lherzolite with 42% MgO.

representative for primary abyssal magma (see Table 1).

Figure 2 shows the compositional relationship between the lherzolitic trend and primary abyssal tholeiite. The primary composition is more or less situated on the extrapolated trend line for most oxides, the exception being SiO_2 . The silica content of the lherzolites displays the smallest relative variation of the oxides, the variation being 45.6–43.3% between 39 and 47% MgO⁴. The trend for SiO_2 is therefore, less well defined than for the other oxides. Although the deviation in the silica content is greater than desirable, because the confidence intervals overlap, the agreement is considered acceptable. The approximate position of the primary abyssal on the extrapolated trend line suggests that the trend may be related to the generation of primary abyssal tholeiite. The present evaluation does not confirm that such a relationship exists, as several tholeiitic compositions fall within the confidence interval of the extrapolated trend line. However, an origin of the lherzolitic trend by primary magma generation is consistent with the composition of primary abyssal tholeiite. The compositions of lherzolite show a variation of 39–47% MgO and this range of compositions is considered to have its origin in the accumulation processes occurring in the mantle. During the accumulation of primary magma some regions of the mantle will be depleted in magma while other regions must

be enriched, as the magma must accumulate before it is erupted. We suggest, therefore, that the trend of the spinel lherzolites originated by a partial melting process which produced primary abyssal tholeiite.

must be situated somewhere on the trend line, either within the composition range of lherzolite, or on the extrapolated trend line between the composition of primary abyssal tholeiite and lherzolite. There is no geochemical or experimental evidence to test these two possibilities. Evidence about possible changes in composition of the mantle since its formation can be obtained by considering the total amount of material it has given up to the oceanic and continental crusts. A model composition of the primitive mantle before continent formation may be approximated by adding, in correct proportions, the compositions of the crusts to the average composition of the present mantle. Assuming that the upper 700 km of the mantle is the depleted part and by using estimated volumes and appropriate densities for the oceanic and continental crusts⁶ the masses are estimated as follows: continental crust, 2.249×10^{19} tons; oceanic crust, 9.223×10^{18} tons; and mantle, 1.022×10^{21} tons. The mass of the continents is thus only 2.1% of the mass of the mantle. The composition used for the mantle is the average composition of spinel lherzolite shown in Table 1. The composition of the continental crust has been estimated by Taylor¹³ (see Table 1) while the composition used for the oceanic crust is that of Elthon¹¹. Thus a model composition of the primitive mantle can be calculated, as shown in Table 1. This composition should be compared with the average composition of lherzolites. The two compositions are evidently very similar, and the formation of the continental crust can only marginally have affected the average composition of the mantle. There is, therefore, no reason to consider compositions outside the composition range of lherzolites as potential mantle compositions, as far as major elements are concerned. The situation for trace elements is different as isotope data for Rb/Sr and Th/U clearly show that the mantle must have been depleted in the incompatible elements^{14,15}, and the concentration of these elements in the lherzolite nodules must consequently differ from those of the primitive mantle. This relationship is also suggested from the present results as both Na_2O and K_2O occur in larger amounts in the calculated primitive mantle than in the average for lherzolites.

We suggest that the variation in major element composition of lherzolite is consistent with a major element composition of the mantle similar to that of lherzolite. A lherzolitic composition with 42.2% MgO is proposed for the average composition of the depleted upper mantle, while the calculated composition with 41.18% MgO is representative of the major element composition of the primitive mantle (Table 1).

Received 9 October 1979; 20 March 1980.

1. Bickle, M. J., Hawkesworth, C. J., Martin, A., Nisbet, E. G. & O'Nions, R. K. *Nature* **263**, 577–580 (1976).
2. DePaolo, D. J. & Wasserburg, G. J. *Geophys. Res. Lett.* **3**, 249–252 (1976).
3. Frey, F. A. & Green, D. N. *Geochim. cosmochim. Acta* **38**, 1023–1059 (1974).
4. Maaløe, S. & Aoki, K. *Contr. Miner.* **63**, 161–173 (1977).
5. Yoder, H. S. *Generation of Basaltic Magma* (National Academy of Science, Washington, DC, 1976).
6. Ronov, A. B. & Yaroslavsky, A. A. *The Earth's Crust and Upper Mantle* (ed. Hurley, P. M.) 45–66 (MIT Press, Massachusetts, 1966).
7. Langmuir, C. H., Bender, J. F., Bence, A. E. & Hanson, G. N. *Earth. planet. Sci. Lett.* **36**, 133–156 (1977).
8. Maaløe, S. *Lithos* **12**, 59–72 (1979).
9. Flower, M. F. J., Robinson, P. T., Schmincke, H. U. & Ohnmacht, W. *Init. Rep. DSDP* **37**, 653–679 (1977).
10. Malpas, J. *Phil. Trans. R. Soc. A238*, 527–540 (1978).
11. Elthon, D. *Nature* **278**, 514–518 (1979).
12. Christensen, N. I. & Salisbury, M. H. *Rev. Geophys. Space Phys.* **13**, 57–86 (1975).
13. Taylor, S. R. *Geochim. cosmochim. Acta* **28**, 1273–1285 (1964).
14. Hofman, A. W. & Hart, S. R. *Earth planet. Sci. Lett.* **38**, 44–62 (1978).
15. O'Nions, R. K., Carter, S. R., Evensen, N. M. & Hamilton, P. J. *A. Rev. Earth planet. Sci.* **7**, 11–38 (1979).
16. Kushiro, I., Shimizu, N. & Nakamura, Y. *Earth planet. Sci. Lett.* **14**, 19–25 (1972).

Carbonate production by algae *Halimeda*, *Penicillus* and *Padina*

Gerold Wefer

Universität Kiel, Geologisches Institut, 2300 Kiel, FRG

The Codiacean green algae *Halimeda* and *Penicillus* and the brown alga *Padina* are important producers of both calcium carbonate and organic matter in shallow water tropical and subtropical areas^{1,2}. Estimates of algal contribution to shallow water carbonate deposition range from 0 to 61%³⁻⁵. However, direct observations on algal carbonate production are very rare. Available data include short-term measurements of calcium and carbon uptake⁶⁻⁹, observations of growth of aquarium specimens¹⁰ and periodic observations of death rate for a year at fixed stations¹¹. I report here on *in situ* measurements taken in Harrington Sound, Bermuda, a shallow subtropical lagoon. Production rates were ~50 (*Halimeda incrassata*), 30 (*Penicillus capitatus*) and 240 *Padina sanctae-crucis* g m⁻² yr⁻¹ calcium carbonate. The measured growth rates suggest that the algae renew their standing stock approximately once every month (*Halimeda* and *Padina*) or once every one and a half months (*Penicillus*) during their growing season.

Alizarin Red-S was used as a time marker to allow measurement of incremental growth of calcareous algae. Algae were covered by a clear plastic tent¹² into which the stain was injected as a concentrated seawater solution. The alizarin was deposited by the calcifying algae while they were thus covered. The stain remains fixed as a visible red band in the skeleton as growth continues after removal of the cover, thus defining a 'time line' in the skeleton (see Fig. 1). All skeletal carbonate deposited after staining seems to be above this time line. The *in situ* measurements were taken during August and September 1978 in the shallow sandy zone in Harrington Sound. The sediment in this zone is predominantly sand with little admixture of coarser or finer material. The coarse and medium sand fractions consist of rock and pelecypod fragments, aragonitic and calcitic algae, serpulid tubes and peneroplid Foraminifera¹³.

For *Penicillus* and *Padina* the distance from the top of the alizarin band to the tip of the plant was measured (see Fig. 1). This increment divided by the number of days between staining and sampling gives the average daily growth. For *Halimeda* the number of segments after staining and the sum of all segments were counted, and a daily growth rate was calculated from the ratio of segments after staining over the sum of all segments present. Several different plant sizes were sampled at any one time, minimizing effects of any nonlinear growth with size.

Altogether, 155 individual plants of *Halimeda incrassata* were sampled 3, 5, 9, 14, 21, 24 and 27 days after staining near Rabbit Island in 2-m water depth. Twelve plants did not grow during staining, 39 recommenced growth within the next 27 days after staining and the remainder continued to grow after the staining event. The daily addition of new segments was ~3% of those present (see Fig 2a). Hence, assuming a steady-state population, the plants should renew their standing stock after about 1 month. The plants can produce up to one segment per day per branch (active growing tips). Plants with several growing branches can produce several segments per day.

Seasonal observations of standing stock made by me during autumn 1977 and spring 1978 and those by other workers^{14,15} indicate that in Bermuda *H. incrassata* grows almost exclusively during summer and autumn, when temperatures are above about 20 °C. Temperatures fall below 20 °C between December and April, reaching a minimum of 14 °C in January¹⁶. During the winter months the standing stock is very low and the plants that do remain become overgrown with other

organisms. For a growing season of 7 months a year, a turnover rate of 7 times per year is expected, if growth is steady over this period. Previous growth rate estimates range from 1 season (6 months)¹⁷ to 1 day¹⁸. The findings of Colinvaux *et al.*¹⁰ are the most similar to those presented here.

A total of 81 specimens of *P. capitatus* were sampled 1, 2, 6, 9, 13 and 20 days after staining from Old Shoals in 8-m water depth. Twelve did not grow during staining, 10 recommenced growth within the next 20 days and the remainder continued to grow after the staining event. The daily average growth was 1.04 mm per day (see Fig. 2b). An average size of 53.3 mm was determined for the *Penicillus* population within the tent. Some 50 days, therefore, would seem to be required for average size to be attained. I found no correlation between plant height and growth, although fast growth (up to 5 mm per day) was observed only in the case of young plants during initial build up of the stalk.

My observations with those of von Bodungen¹⁵ and Bernatowicz¹⁴ indicate that in Harrington Sound the growth period for *Penicillus*, like *Halimeda*, lasts only 7 months. Between December and April the standing stock is very low. The plants that do remain become overgrown with other organisms, especially bryozoans. The overgrowth extends to the upper tips and thus the plant seems to be dormant. If steady-state growth for the summer and autumn months is assumed the turnover rate for *P. capitatus* would be about 4 times per year. These observations agree with the suggestion of Rezak¹⁷, that natural occurrences of *Penicillus* indicate a life span for this plant of 30–60 days. Stockmann *et al.*¹¹ determined an average life span of ~40 days for *Penicillus* in Florida Bay and 60 days for the inner reef tract on the South Florida Shelf. Colinvaux *et al.*¹⁰ observed that *Penicillus* shows sporadic rapid growth lasting 1 or perhaps 2 weeks, followed by long periods of non-growth. Maturity seems to be reached within about 1 month, according to their data.

A total of 125 fronds of *P. sanctae-crucis* were sampled 1, 2, 4, 6, 8 and 11 days after staining near Rabbit Island in 1.5 m water depth. Eight of them did not grow during staining, three recommenced growth within the next 11 days and the remainder continued to grow after the staining event. The average linear growth on fronds was 0.81 mm per day (see Fig. 2c). For the whole *Padina* community inside the tent an average plant height of 20.4 mm was determined. Hence, the algal population would

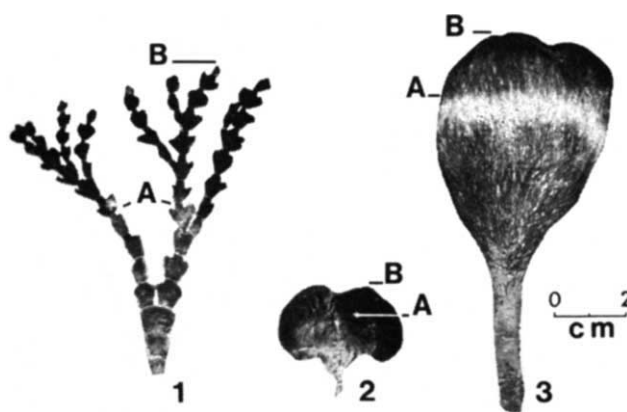


Fig. 1 *Halimeda incrassata* (1), *Padina sanctae-crucis* (2) and *Penicillus capitatus* (3). Photographs of stained plants 7 days (*Halimeda*), 11 days (*Padina*) and 15 days (*Penicillus*) after the staining event. A, Top of the Alizarin line; A–B, increment between staining and sampling. The plants were covered by a tent with a volume of ~2 m³. The enclosed bottom area was 3 m². The concentration within the tents was about 1.5 mg Alizarin Red-S per 1 seawater. Tents remained in place for 1 day. Algal colonies were sampled in 1-day to 1-week intervals.

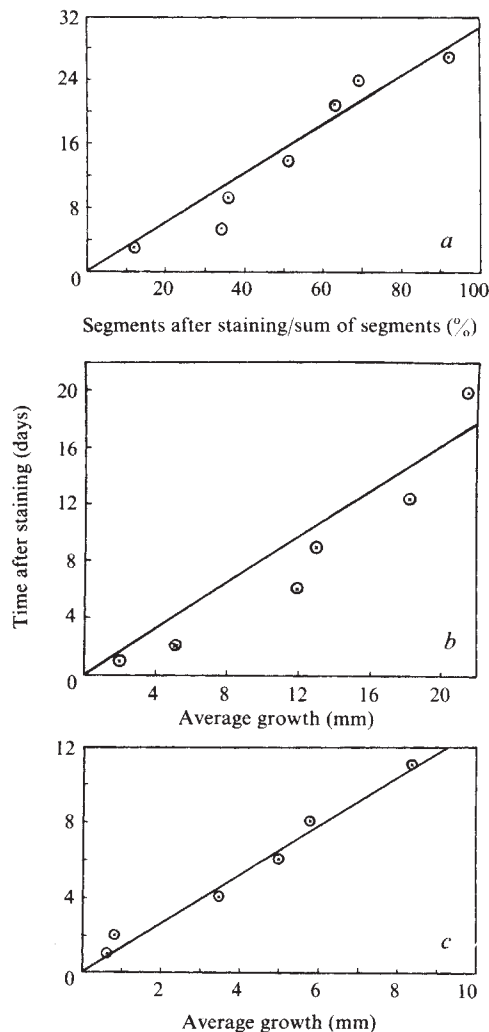


Fig. 2 *a*, *Halimeda incrassata*. Growth rate expressed as ratio between the new segments after staining and the total number of algal segments for the sampling dates. Each point represents the average of a representative number (10–20) of algae of the whole standing stock. The average (s) of the standard deviations of the data for each point is 23. Daily average increment of the standing stock is 3%. *b*, *Penicillus capitatus*. Average growth rates for several sampling dates. ($s = 6.2$) Daily average growth for the sampled population is 1.04 mm per day. *c*, *Padina sanctae-crucis*. Average growth rates for several sampling dates. ($s = 1.1$) Daily average growth for the sample population is 0.81 mm per day.

need ~25 days to build up the standing stock sampled in the area within the tent.

During March 1978 *P. sanctae-crucis* was rare on the cliffs and rocks of the shallow area of Harrington Sound. With an increase of temperature from 18 to 20 °C in April 1978, *Padina* started to grow. Apparently, like *Halimeda* and *Penicillus*, *Padina* has its growth season during the warm period of the year. Assuming steady-state growth of the population for the summer and autumn months, the standing stock would be renewed about 8 times per year, somewhat more often than that of *Halimeda* and *Penicillus*. No data on growth are available in the literature. However, it is clear that *Padina* cannot be regarded as only a minor carbonate contributor as assumed by Milliman². On the cliffs and submarine rocks of Harrington Sound (Bermuda), at a depth of 1–3 m a standing stock was measured of up to 350 g dry wt m⁻². With a carbonate content of 38% of dry weight¹⁹, and at a possible turnover rate of up to 8 times per year the annual carbonate production could be as high as 1,060 g m⁻² (the average is 240 g m⁻² yr⁻¹). This figure is actually above the range observed for other calcareous algae in

Harrington Sound. For *Halimeda*, the carbonate production is 50 g m⁻² yr⁻¹ based on the assumed turnover rate of 7 times per year and an average standing stock of 6.7 g CaCO₃ per m² found at 31 stations arbitrarily selected in 1–4 m water depth. For *Penicillus* the carbonate production is 30 g m⁻² yr⁻¹ based on the assumed turnover rate of 4 times per year and an estimated average standing stock of 5.6 g CaCO₃ m⁻² (at the 31 stations). Another important algal carbonate contributor is *Halimeda monile*, which is only common in areas equal to or shallower than 3 m and may contribute as much carbonate as *H. incrassata*. Other green algae like *Udotea*, *Cymopolia* and *Rhipocephalus* are mostly solitary, rare or absent in Harrington Sound and are not important as carbonate producers. The red alga *Amphiroa fragilissima* abounds on rocky substrates and may be a major carbonate producer here.

The total carbonate production by algae in the sandy area of Harrington Sound is nearly 100 g CaCO₃ per m² per yr and on the rocky substrate nearly 500 g CaCO₃ per m² per yr, which is about half of the carbonate deposition in the shallow water area of the Sound. The carbonate deposition is assessed from the bulk composition of dated sediments collected by cores. The various algal species are restricted to certain types of bottom consistency, and their relative contribution will, therefore, depend on locality. For example, *Amphiroa* and *Padina* are the dominant carbonate producers on rocky substrate and almost absent in sandy areas, while *Halimeda* and *Penicillus* are the major carbonate contributors in the sand covered regions of Harrington Sound.

I thank V. Smetacek, A. C. Neumann and W. H. Berger for helpful criticisms. Supported by Deutsche Forschungsgemeinschaft.

Received 27 December 1979; accepted 14 March 1980.

1. Bathurst, R. G. C. *Carbonate Sediments and their Diagenesis*, 1–620 (Elsevier, Amsterdam, 1971).
2. Milliman, J. D. *Marine Carbonates*, 1–375 (Springer, Berlin, 1974).
3. Milliman, J. D. *U.S. Geol. Surv. Prof. Pap.* 529–7 (1972).
4. Ginsburg, R. N. *Am. Ass. Petrol. Geol. Bull.* **40**, 2384–2427 (1956).
5. Neumann, A. C. & Land, L. S. *J. Sediment. Petrol.* **45**, 763–786 (1975).
6. Goreau, T. F. *Ann. N.Y. Acad. Sci.* **109**, 127–167 (1963).
7. Stark, K. W., Ginsburg, R. N. & Shinn, E. A. *J. Sediment. Petrol.* **37**, 633–648 (1967).
8. Böhm, E. L. *Mar. Biol.* **47**, 9–14 (1978).
9. Borowitzka, M. A. & Larkum, A. W. D. *J. exp. Bot.* **27**, 864–878 (1976).
10. Colinvaux, L. H., Wilbur, K. M. & Watabe, N. *J. Phycol.* **7**, 69–78 (1965).
11. Stockmann, K. W., Ginsburg, R. N. & Shinn, E. A. *J. Sediment. Petrol.* **37**, 633–648 (1967).
12. Dustan, P. *Mar. Biol.* **33**, 101–107 (1975).
13. Neumann, A. C. *Bull. mar. Sci.* **15**, 987–1035 (1965).
14. Bernatowicz, A. J. *Pap. Mich. Acad. Sci. Arts Lett.* **36**, 3–8 (1952).
15. von Bodungen, B. *Bermuda Biological Station, Spec. Publ.* (in the press).
16. Barnes, J. A. & von Bodungen, B. *Bermuda Biological Station, Spec. Publ.* **17**, 1–190 (1978).
17. Rezak, R. in *Geology of Calcareous Algae* (eds Ginsburg, R. N., Rezak, R. & Wray, J. L.) 3.1–3.8 (The Comparative Sedimentology Laboratory, Miami, 1971).
18. Goreau, T. F. in *The Biology of Hydra*, 269–285 (University of Miami Press, 1961).
19. Böhm, E. L. & Schramm, W. in *Das Harrington Sound Projekt* (eds Wefer, G. & Hempel, G.) 20–27 (Sonderforschungsbereich 95, Kiel University, 1977).

Effects of density on monospecific stands of marine algae

David R. Schiel & J. H. Choat

Department of Zoology and Leigh Marine Research Laboratory, University of Auckland, New Zealand

Current ideas on intraspecific competition between plants in monospecific stands stem from work on terrestrial floras. The central tenet is that higher plant density exacts a price from the individuals in a stand: (1) total yield (biomass) per area reaches a constant value at high density (the law of constant final yield^{1–5}); (2) the weight distribution of individuals becomes skewed, with few large and many small plants^{1,5–8}; (3) survival rates may be reduced when plants are sown at higher densities^{1,9}.

The effects of density seem to be remarkably similar for primitive as well as higher plants^{1,2,9-11} and are often summarized by the 3/2 thinning law^{1,2,5-7,9-11}, which indicates that if a stand of plants is sampled through time, the relationship between log weight and log density is a line of -1.5 slope. The basis of this is that mortality is density dependent, so that density is related to time through the thinning process due to crowding. We report here that these effects do not obtain for two species of subtidal marine algae. For the characters which we examined—total yield, plant length, plant dry weight and dry weight of reproductive parts—algae fared better at higher densities. We conclude that there is a difference between marine and terrestrial plants in their responses to density, and we suggest a difference in herbivory due to arthropod-plant associations in each system.

Our investigations were prompted by observations that algal plants of both stunted and lush forms, characters usually ascribed to different degrees of wave exposure¹²⁻¹⁴, occur metres apart on the same shore in similar conditions of exposure. We hypothesized plant density to be an important factor. The questions we posed were, what effects does density have on total yield per given area, on plant size, weight, reproductive output and survival? We examined two species of algae which are quite different in form and life history. The laminarian alga, *Ecklonia radiata*, has a single erect stipe, with a primary and numerous secondary laminae; it has a life span of about 7 yr and occurs in typically dense stands in the northern New Zealand subtidal zone. The furoid alga, *Sargassum sinclairii*, is a facultative annual bearing discrete reproductive parts and many small laminae on primary and lateral branches; it is very abundant in shallow depths. For each species, we examined monospecific, even-aged stands (1-yr old *Sargassum*, 3-yr old *Ecklonia*) from a shallow (2-6 m) open coast, tumbled boulder area. In addition to using naturally occurring populations, *Sargassum* germlings were settled at four different densities ($\bar{x} = 4.4, 12.8, 20.3, 29.7$ per cm^2) on asbestos-cement plates and transferred to the field for observations on survival. In all cases the major herbivores—sea urchins and gastropods—were excluded.

Table 1 Size frequency data for *Sargassum* and *Ecklonia*

<i>Sargassum sinclairii</i>	<i>Ecklonia radiata</i>
Low density (1-50 plants m^{-2})	Low density (1-6 plants m^{-2})
Length: $P = 0.05$, positive skew	Length: NS
Dry weight: $P = 0.01$, positive skew	Dry weight: NS
Dry weight of reproductive parts: $P = 0.01$, positive skew	
High density (160-220 m^{-2})	High density (30-45 m^{-2})
Length: NS	Length: NS
Dry weight: NS	Dry weight: NS
Dry weight of reproductive parts: NS	

G-tests for skewness of distribution²¹ were used to test the frequency distributions of length, plant dry weight (*Sargassum* and *Ecklonia*) and dry weight of reproductive parts (*Sargassum*) at low and high densities. NS, Non-significant result, normal distribution.

Figure 1a,b shows that total yield did not reach a constant value up to the highest plant densities found. The relationship of plant length to density was a line of positive slope for both species (Fig. 1c,d). This is not necessarily unusual because terrestrial plants in dense stands may become etiolated^{1,15}. However, the slopes were also positive for regressions of individual plant dry weight on density (Fig. 1e,f) and for the dry weight of reproductive parts on density for *Sargassum* (Fig. 1g). The largest plants by any measure generally occurred at higher densities. These relationships do not apply to all algae; a constant yield at different densities is recorded for the red alga *Porphyra*¹⁶. However, there is a suggestion that they may apply to at least one Japanese laminarian¹⁷.

Table 1 summarizes the size frequency data. For *Sargassum*, high-density plants were normally distributed for all measures whereas at low density distributions were skewed, with most plants in the smaller size classes. *Ecklonia* at both low and high densities showed normal distributions for all measures. Few algal plants were trapped under the canopy at high density, in contrast to terrestrial systems which tend to have many smaller plants suppressed below the canopy of a few larger individuals in crowded stands^{1,5-8}.

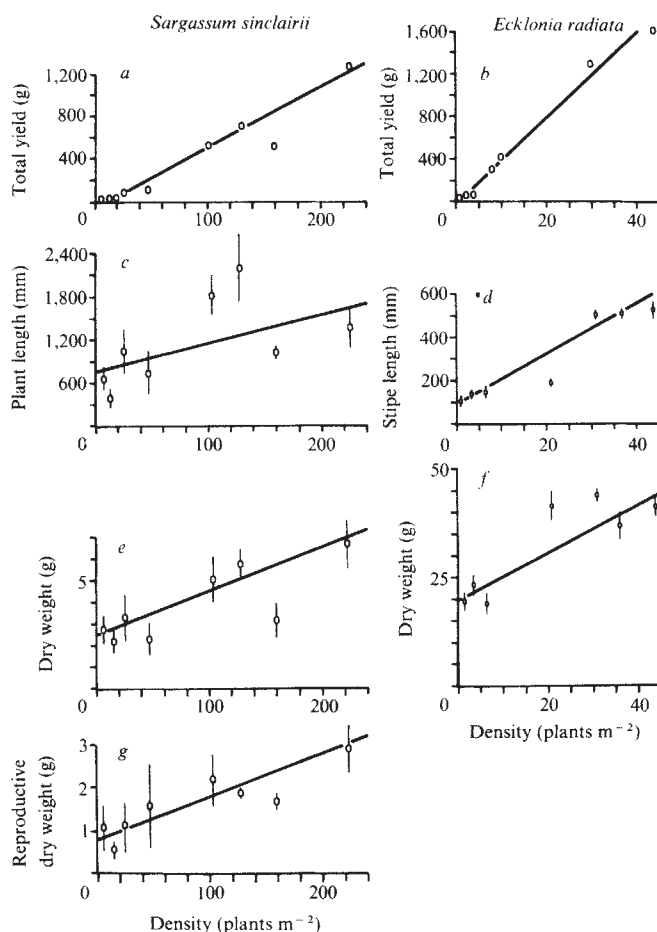


Fig. 1 Graphs relating various measures and density for *Sargassum sinclairii* and *Ecklonia radiata*. Regression equations, the probability of $\beta = 0$ and r^2 values are respectively: (a) $y = 5.66x - 70.34$, $P < 0.001$, $r^2 = 0.88$; (b) $y = 39.32x - 1.65$, $P < 0.001$, $r^2 = 0.99$; (c) $y = 3.80x + 757.33$, $P < 0.007$, $r^2 = 0.16$; (d) $y = 11.68x + 87.96$, $P < 0.001$, $r^2 = 0.66$; (e) $y = 0.02x + 2.46$, $P < 0.001$, $r^2 = 0.25$; (f) $y = 0.54x + 19.84$, $P < 0.001$, $r^2 = 0.59$; (g) $y = 0.01x + 0.81$, $P < 0.003$, $r^2 = 0.19$. These relationships were calculated using least squares linear regression. Each y value represents the measure of a single plant, except for the total yield against density regressions, where total dry weight per m^2 was used. For clarity of presentation, means ± 1 s.e. are plotted, but regressions were calculated using all x, y values. We realize that regression analysis may be inappropriate here and that other models may better describe these relationships. However, as the variance is high and transformations produced no significant change in r^2 , we have used a simple linear model as a general description.

An examination of survival of the artificially settled *Sargassum* germlings showed that the rate of decrease in numbers through time was the same at four different initial densities (number of plants surviving regressed with time; covariance, $n = 4$, $F = 2.61$, $P = 0.10$). However, mortality of naturally occurring *Ecklonia* sporophytes was found to be density

dependent, but not markedly so (proportion of plants surviving after 1 yr regressed on initial densities: $y = -0.0054x + 0.65$, $P(\beta=0) < 0.001$, $r^2 = 0.67$). This slightly higher proportional mortality at increased density is more than balanced, however, by greater numbers of plants per unit area and larger individual size.

Because of these density responses, the 3/2 thinning law is unlikely to apply to these marine algae. It can be argued that higher densities might produce some of the effects documented for terrestrial plants, and certainly there must be an upper limit to plant density and biomass per unit area. However, at the highest densities we found, even though thinning does occur, that there is no consequent compensation in size by other members of the stand.

We consider these density responses to be primarily the result of two factors: (1) both *Sargassum* and *Ecklonia* show faster growth for high density plants (J.H.C. and D.R.S., in preparation), and (2) there may be some protection from physical battering in the turbulent, shallow coast environment for high-density individuals which may be cushioned in their swaying by others in the stand. Our observations of a range of plant forms in the same area but at different densities suggest that the water movement intensity, which is said to influence morphological differences in some laminarians¹²⁻¹⁴, may be altered by high density.

We also observe a noteworthy difference between terrestrial and marine systems in the nature of their plant-arthropod associations. Adult terrestrial plants often become infested with herbivorous insects, which can cause severe defoliation^{18,19}; herbivore loads may be greatest in monospecific stands²⁰. There is no equivalent form of herbivory on large marine plants. Although the arthropod load may be high (we have found up to 1,700 copepods, amphipods and isopods per plant), these animals do not visibly damage the algal fronds and there seems to be no high density debit due to herbivory.

Large, dense, predominantly single-species brown algal stands are a common feature of most boreal and temperate shores. Our knowledge of competition based on density stress in terrestrial plants seems to be inappropriate to deal with these patterns of algal abundance. If these density responses hold generally, they suggest two fruitful areas of comparison with terrestrial floras: (1) the relationship between plant density and herbivory, particularly plant-arthropod associations, and (2) the consequences of inter- and intraspecific competition in stands of large brown marine algae. In addition to clarifying the relationships of the dominant organisms in algal zones, there could also be useful implications for the mariculture of seaweeds.

We thank J. Ogden, B. McArdle, R. Black and M. Hawkes for criticizing the manuscript, and the Auckland University Grants Committee and the Leigh Marine Laboratory for providing funds and facilities.

Effects of methylated xanthines on mammalian cells treated with bifunctional alkylating agents

John P. Murnane, John E. Byfield, John F. Ward & Paula Calabro-Jones

Division of Radiation Oncology, University Hospital, and University of California San Diego Cancer Center, 225 Dickinson Street, San Diego, California 92103

Caffeine has been previously reported to enhance the lethal potential of many DNA-damaging agents in rodent cells¹⁻⁵. This effect has most commonly been ascribed to the binding of caffeine to single-stranded DNA⁶, and the resulting inhibition of post-replication repair⁷⁻¹⁰, which is associated with the synthesis of abnormally small nascent DNA fragments^{7,11-13}. However, certain aspects of this theory remain unclear: (1) why does the addition of caffeine to damaged cells elevate the level of DNA synthesis when it supposedly blocks post-replication repair^{10,14}, and (2) as pointed out by Cleaver¹⁵, why does caffeine continue to exert its synergistic lethal effects until completion of the S phase^{16,17}, even though the size of newly synthesized DNA seems normal much earlier¹⁸⁻²⁰? The present studies with nitrogen mustard (HN₂) fail to demonstrate any effect of non-lethal concentrations of methylated xanthines (MXs) on removal of DNA damage or post-replication repair in conditions producing synergistic lethal effects. We demonstrate an influence by MXs on initiation of DNA synthesis in damaged replicons, and propose that this effect is primarily responsible for the synergistic lethal properties of these drugs.

The effects of caffeine on damaged human cells are less consistent than those on rodent cells^{16,21-24}. Figure 1 shows the survival curves obtained when human HT-29 cells (ref. 25, human karyotype confirmed by K. Benirschke) were exposed to increasing concentrations of HN₂ with or without a non-lethal concentration (1 mM) of caffeine. As shown by the reduction in the shoulder width of the HN₂ survival curve, human HT-29 cells seem to be highly sensitive to the synergistic lethal effects of caffeine in combination with HN₂. Similar results were obtained when mouse cells were used, melphalan or phosphoramide mustard (PM) were substituted in place of HN₂ or when caffeine was replaced by theobromine or pentoxiphylline.

DNA cross-links are thought to be responsible for the relatively high toxicity of bifunctional alkylating agents²⁶. We have used the technique of alkaline elution developed by Kohn *et al.*²⁷ to estimate the appearance and removal of DNA cross-links^{28,29}. As shown in Fig. 2, 1 mM caffeine has no effect on either the maximum (6 h) or residual (24 h) levels of cross-links after a 120- μ M dose of PM in either human (70% lethal) or mouse cells (95% lethal). Caffeine, therefore, has no effect on either cross-link formation or removal, and hence, its synergistic lethal effects are apparently unrelated to DNA repair.

As the synergistic lethal effects of caffeine have been reported to be associated with the inhibition of nascent DNA elongation in rodent cells, we have studied the relationship between these two phenomena in HN₂-treated cultures. Pulse-chase techniques have been combined with alkaline elution as done previously to demonstrate elongation of newly synthesized DNA³⁰. The results shown in Fig. 3 are for mouse cells; similar results were obtained for human HT-29 cells. Figure 3a shows that ~6 h are required for newly synthesized DNA from untreated C3H-10T_{1/2} cells to elute at the rate of DNA continuously labelled for 24 h. DNA from cells pulsed 0.5 h after removal of a 99% lethal dose of HN₂ (Fig. 3b) requires much longer (>12 h) to reach full size, as reported previously³¹. HN₂-treated DNA seems initially (0 h) to elute more slowly than DNA from control

Received 4 February; accepted 10 March 1980.

- Harper, J. L. *Population Biology of Plants* (Academic, London, 1977).
- Kays, S. & Harper, J. L. *J. Ecol.* **62**, 97-105 (1974).
- Donald, C. M. *Aust. J. agric. Res.* **2**, 355-378 (1951).
- Kira, T., Ogawa, H. & Sakazaki, N. *J. Inst. Polytech. Osaka City Univ.* **D 4**, 1-16 (1953).
- Harper, J. L. & White, J. in *Dynamics of Populations* (eds den Boer, P. J. & Gradwell, G. R.) 41-63 (Centre for Agricultural Publications and Documentation, Wageningen, 1971).
- Mohler, C. L., Marks, P. L. & Sprugel, D. G. *J. Ecol.* **66**, 599-614 (1978).
- Schlesinger, W. H. & Gill, D. S. *Ecology* **59**, 1256-1263 (1978).
- Ogden, J. *Proc. N.Z. ecol. Soc.* **17**, 1-9 (1970).
- Yoda, K., Kira, T., Ogawa, H. & Hozumi, K. *J. Biol. Osaka City Univ.* **14**, 107-129 (1963).
- White, J. & Harper, J. L. *J. Ecol.* **58**, 467-485 (1970).
- Gorham, E. *Nature* **279**, 148-150 (1979).
- Chapman, A. R. O. *Phycologia* **12**, 53-57 (1973).
- Gerard, V. A. & Mann, K. H. *J. Phycol.* **15**, 33-41 (1979).
- Mann, K. H. *J. Fish. Res. Bd Can.* **28**, 778-780 (1971).
- Hodgson, G. L. & Blackman, G. E. *J. exp. Bot.* **7**, 147-165 (1956).
- Yoshida, T. *Bull. Tohoku reg. Fish. Res. Lab.* **32**, 89-94 (1972).
- Kawashima, S. in *Contributions to the Systematics of Benthic Marine Algae of the North Pacific* (eds Abbott, I. A. & Kurogi, M.) 93-108 (Japanese Society of Phycology, Kobe, 1972).
- Feeny, P. *Ecology* **51**, 565-579 (1970).
- Morrow, P. A. & La Marche, V. C. *Jr Science* **201**, 1244-1246 (1978).
- Root, R. B. *Ecol. Monogr.* **43**, 95-120 (1973).
- Sokal, R. R. & Rohlf, F. J. *Biometry* (Freeman, San Francisco, 1969).

cells (Fig. 3b). A delay of 4 h after the removal of the HN_2 before pulse labelling eliminates this effect (data not shown).

Despite the increased lethality due to caffeine in HN_2 -treated cells, 1 mM caffeine did not affect DNA elongation in control or HN_2 -treated cultures in many experiments (Fig. 3), regardless of whether caffeine was added before or after pulse labelling. That this is true for mouse cells as well as human indicates that these results are not simply due to cell type, as alkaline elution studies in the same mouse cell lines demonstrated that caffeine inhibited DNA elongation after UV irradiation (data not shown). As reported by others⁷⁻¹⁰, HN_2 damage, therefore, seems different from UV in that HN_2 -treated rodent cells sensitized with caffeine demonstrate no inhibition of nascent DNA elongation. This failure of caffeine to inhibit elongation is not simply an artefact of the alkaline elution technique, as these results have been confirmed for HT-29 cells using alkaline sucrose gradients

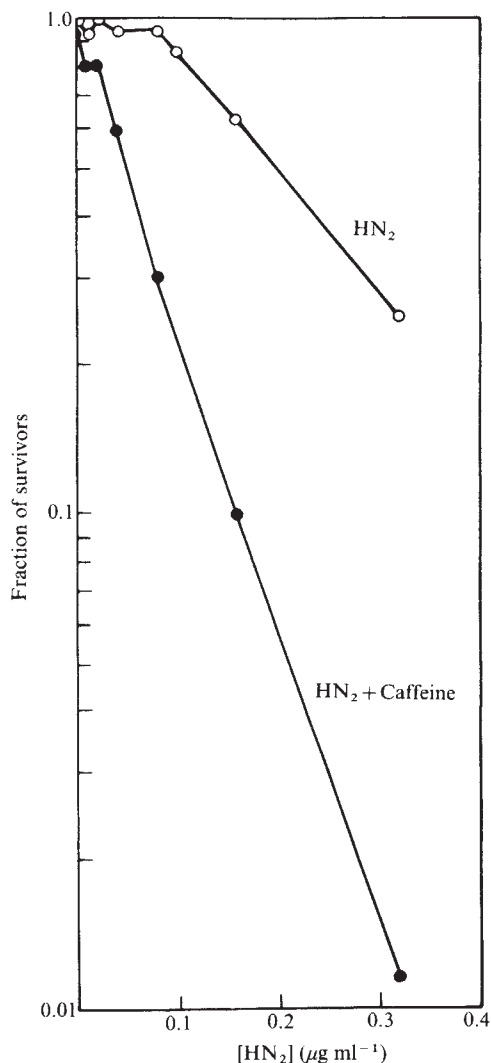


Fig. 1 Effect of HN_2 on survival of HT-29 cells with (●) and without (○) 1 mM caffeine (Sigma). Human HT-29 cells (400) were plated onto 60-mm Lux tissue culture dishes in 2 ml of Eagle's basal medium (BME, Gibco) with 10% fetal calf serum (FCS, Gibco). After 1 h at 37 °C in a humidified incubator with a 5% CO_2 atmosphere, 1 ml BME plus 10% FCS containing the various concentrations of HN_2 were added to the appropriate plates. Cells were reincubated for 2 h at 37 °C, after which the medium in all plates was removed. BME (2 ml) plus 10% FCS containing the 1 mM caffeine were then added to both HN_2 -treated and untreated cells (in triplicate). After a further 72 h incubation, the medium was removed and replaced with fresh BME plus 10% FCS without caffeine. The cells were then incubated for 4 more days before fixation with methanol, Giemsa staining and counting of visible colonies.

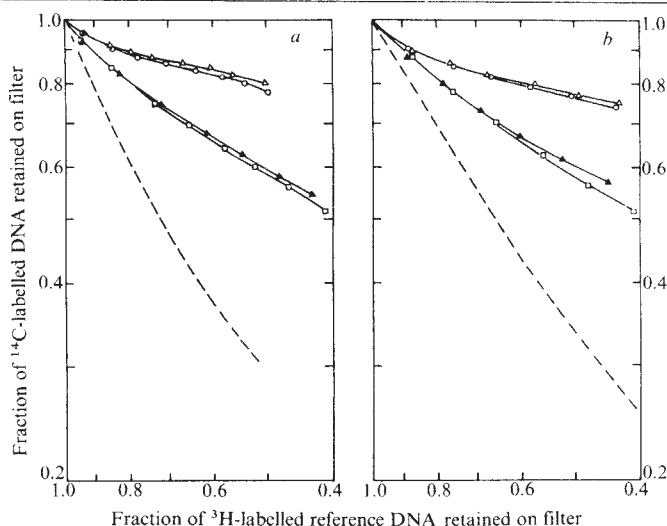


Fig. 2 Alkaline elution profiles of DNA from HT-29 (a) and C3H-10T½ (b) cells treated with 120 μM PM in BME plus 10% FCS for 2 h, rinsed and reincubated in BME plus 10% FCS for 6 (○) or 24 (□) hours; or 120 μM PM plus 1 mM caffeine in BME plus 10% FCS incubated for 2 h, rinsed and reincubated in 1 mM caffeine in BME plus 10% FCS for 6 (△) or 24 (▲) h. Alkaline elution was carried out as previously described^{27,28}. Cells grown in BME plus 10% FCS were labelled for 24 h before drug treatment with 5 ml of either [$6\text{-}^3\text{H}$]thymidine (28.5 Ci mmol^{-1} , 0.1 $\mu\text{Ci ml}^{-1}$; NEN) for reference cells, or [$2\text{-}^{14}\text{C}$]thymidine (52.4 mCi mmol^{-1} , 0.01 $\mu\text{Ci ml}^{-1}$; NEN) for test cells. Before elution, ^3H -labelled reference cells were irradiated with 300 R and test cells with 600 R of X rays (Picker orthovoltage, 280 keV, 1 mm copper filter) at 0 °C, and 5×10^5 cells of each added to the elution filter. Dashed lines represent the elution profiles of control cells with or without 1 mM caffeine for 24 h. Control cells were labelled for 24 h similarly to test cells, but with a medium change 2 h before elution.

similar to those shown in Fig. 4. Thus, inhibition of DNA elongation and increased lethality do not seem to be directly related in this system, and some mechanism other than inhibition of post-replication repair is necessary to explain the effects of caffeine on cell survival.

During these experiments we found that addition of caffeine to HN_2 -treated cells reversed the inhibition of DNA synthesis due to DNA damage, as previously reported by others using caffeine in combination with UV¹⁰, *cis*-platinum II (ref. 14) and X rays³². We have since established a correlation between the synergistic lethal effects of various MX analogues and their ability to produce increased DNA synthesis in HN_2 -treated cells (in preparation). We have, therefore, investigated this phenomenon further.

The decreased rate of DNA synthesis in damaged cells has been proposed to be a result of the inhibition of replicon initiation in DNA domains containing damaged sites. X rays^{31,33,34}, bromouracil photolysis³⁵ and methyl methanesulphonate³⁶ inhibit replicon initiation, and we have observed a similar response to HN_2 damage (Fig. 4). We have determined whether caffeine increases damaged DNA synthesis by influencing replicon initiation. Using an alkaline sucrose procedure which demonstrates replicon size distribution and elongation³³, we found that the application of 1 mM caffeine to undamaged cells reduces the size of the DNA that is normally synthesized during a short ^3H -thymidine pulse (Fig. 4). This reduction in the quantity of larger DNA fragments has previously been reported to be due to an inhibition of elongation of nascent DNA⁹. However, the absence in undamaged cells of any effect of 1 mM caffeine on colony formation (Fig. 1), the rate of DNA synthesis (in preparation) or the time required for full elongation of DNA (ref. 9 and Fig. 3a), lead us to concur with

Lehman³⁷ that an increased number of replicon initiation sites is a more likely explanation.

In contrast to the caffeine-induced reduction in the size of newly synthesized DNA in undamaged cells, the addition of 1mM caffeine to HN₂-treated cells induces the reappearance of a small-size class of DNA previously eliminated by the inhibition of replicon initiation due to DNA damage. The influence of MXs on replicon initiation is therefore capable of interfering with the regulation of DNA synthesis on damaged DNA templates. This increased initiation of DNA synthesis in damaged replicons would greatly increase the number of replication forks encountering damaged sites, and could account for the inhibitory effects of caffeine on nascent DNA elongation in rodent cells through saturation of the post-replication repair system. The variability of the effects of MXs on DNA elongation could therefore be due to differences in mechanisms used for replicative bypass of DNA damage among various cells and with various types of lesions.

The exact mechanism by which MXs influence control of replicon initiation is unclear. The inhibition of cyclic nucleotide phosphodiesterase by MXs does not seem to be involved, as drugs influencing cyclic AMP or cyclic GMP do not react similarly to caffeine³². Chromatin conformation has been postulated to regulate both normal DNA synthesis³⁹ and inhibition of damaged replicon initiation³³, suggesting that MX influences the availability of initiation sites by alteration of chromatin structure.

Reports of increased survival in cells held in a non-dividing state following treatment with various DNA-damaging agents^{40,41}, and of a reduced shoulder and D_0 when cells were irradiated with near-UV light during replication of DNA

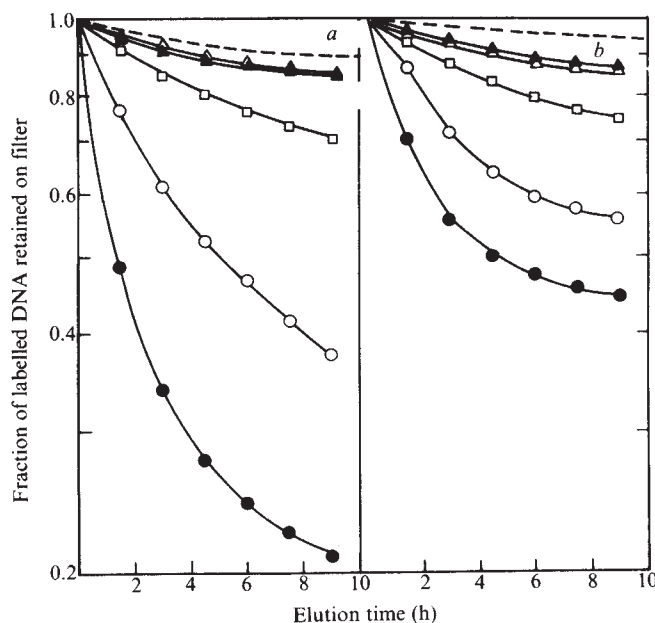


Fig. 3 Alkaline elution profiles of C3H-10T_{1/2} cells DNA grown in BME plus 10% FCS and pulse labelled with 5 ml of [6-³H]thymidine (28.5 Ci mmol⁻¹, 0.4 μ Ci ml⁻¹) for 20 min: *a*, without prior HN₂ treatment and chased for 0 h (●), 2 h (○), 4 h (□) and 6 h with (▲) or without (△) 1 mM caffeine; *b*, pulsed 0.5 h after changing media from a 2-h incubation with 0.33 μ g ml⁻¹ HN₂ and incubating further in cold media for 0 h (●), 4 h (○), 8 h (□), 12 h with (▲) or without (△) 1 mM caffeine. Untreated cells continually labelled for 24 h are shown by a dashed line in *a* and pulse-labelled cells treated with HN₂ and chased for 30 h are shown by dashed lines in *b*. Alkaline elution was carried out as previously described³⁰. 10⁶ cells were applied to each elution filter after being scraped into cold phosphate-buffered saline (PBS) and stored on ice. The addition of 10 μ M cold thymidine in the chase media increased the apparent rate of elongation, but did not influence the absence of effect of caffeine.

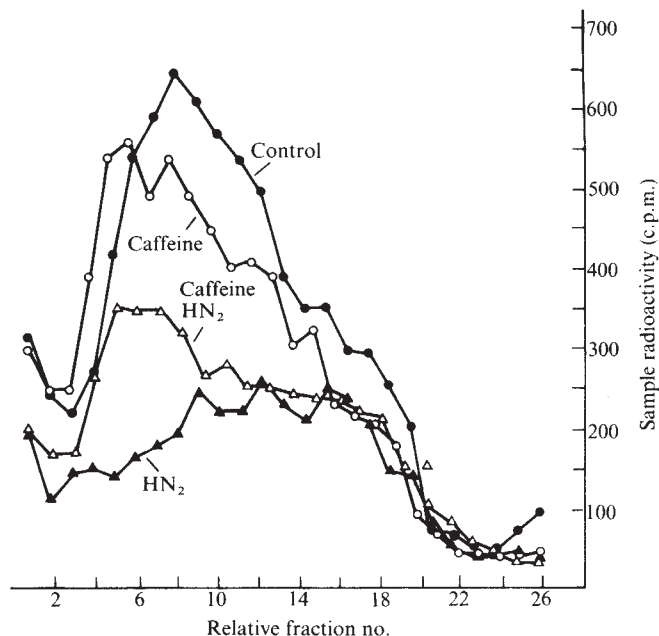


Fig. 4 Alkaline sucrose gradients for HT-29 cells pulse labelled for 10 min with [6-³H]thymidine (28.5 Ci mmol⁻¹, 20 μ Ci ml⁻¹) after HN₂ treatment (▲), HN₂ treatment plus caffeine (△), caffeine alone (○) and control cells (●). Sedimentation is from left to right. After a 2-h treatment with 0.33 μ g ml⁻¹ HN₂, cells were rinsed and incubated for 30 min with or without caffeine, and then given a 10-min pulse of ³H-thymidine. Control cells with or without caffeine were handled similarly but without HN₂. Caffeine-treated cells also received 1mM caffeine during the ³H-thymidine pulse. The treatment of cells and the alkaline sucrose procedure have been previously described³². After the pulse, cells were rinsed in cold PBS containing 10⁻³ M thymidine, scraped off the plates and irradiated with 1,000 R of X rays. 10⁵ cells in 0.5 ml of PBS were then layered onto a 0.5-ml layer of lysis solution on top of a linear 5–20% alkaline sucrose gradient. After 15 min in the dark, the gradients were spun for 3 h at 27 K, and 25 fractions were collected from the bottom. Calf thymus DNA (100 μ g) was added to each fraction, 50% trichloroacetic acid added to a final concentration of 5%, and all samples refrigerated overnight. Samples were filtered on GF/C glass fibre filters (Whatman) using cold 5% trichloroacetic acid. For determination of radioactivity, 0.4 ml of 1 M HCl was added to each filter in scintillation vials, and the vials heated at 70 °C for 1 h. After cooling, 2.5 ml of 0.4 M NaOH and 6 ml Aquasol (NEN) were added and the samples counted in gel phase.

regions containing bromouracil⁴² indicate a correlation between damaged DNA replication and increased lethality. Thus, the inhibition of initiation of DNA synthesis due to damage may promote cell survival, and therefore damage which is 'potentially lethal' may be rendered lethal by inducing damaged replicon initiation.

This work was supported by funds from the University of California San Diego Cancer Center. We thank Megan Cunningham for technical assistance. The stock HT-29 cells were provided by Dr J. Fogh and C3H-10T_{1/2} cells by Dr C. Heidelberger.

Received 17 January; accepted 19 March 1980.

1. Rauth, A. M. *Radiat. Res.* **31**, 121–138 (1967).
2. Rauth, A. M., Barton, C. & Lee, C. P. Y. *Cancer Res.* **30**, 2724–2729 (1970).
3. Walker, I. G. & Reid, B. D. *Mutat. Res.* **12**, 101–104 (1971).
4. Roberts, J. J., Sturrock, J. E. & Ward, K. N. *Mutat. Res.* **26**, 129–143 (1974).
5. Busse, P. M., Bose, S. K., Jones, R. W. & Tolmach, L. J. *Radiat. Res.* **71**, 666–677 (1977).
6. Ts'o, P. O. P. & Lu, P. *Proc. natn. Acad. Sci. U.S.A.* **51**, 17–24 (1964).
7. Cleaver, J. E. & Thomas, G. H. *Biochem. biophys. Res. Commun.* **36**, 203–208 (1969).
8. Fujiwara, Y. *Exptl Cell Res.* **75**, 483–489 (1972).
9. Buhl, S. N. & Regan, J. D. *Biophys. J.* **14**, 519–527 (1974).
10. Lehmann, A. R. & Kirk-Bell, S. *Mutat. Res.* **26**, 73–82 (1974).
11. Rupp, W. D. et al. *Brookhaven natn. Lab. Publ. No.* 50203, 1–11 (1969).
12. Lehmann, A. R. *J. molec. Biol.* **66**, 319–337 (1972).
13. Buhl, S. N., Setlow, R. B. & Regan, J. D. *Int. J. Radiat. Biol.* **22**, 417–424 (1972).
14. Van Den Berg, H. W. & Roberts, J. J. *Chemico-Biol. Interact.* **12**, 375–390 (1976).

15. Cleaver, J. E. *Biochim. biophys. Acta* **516**, 489–516 (1978).
16. Roberts, J. J. & Ward, K. N. *Chemo-Biol. Interact.* **7**, 241–264 (1973).
17. Dorn, J. & Rauth, A. M. *Radiat. Res.* **39**, 207–214 (1969).
18. Meyn, R. E. & Humphrey, R. M. *Biophys. J.* **11**, 295–301 (1971).
19. Lehmann, A. R. *Eur. J. Biochem.* **31**, 438–445 (1972).
20. Buhl, S. N., Setlow, R. B. & Regan, J. D. *Biophys. J.* **13**, 1265–1275 (1973).
21. Wilkinson, R., Kiefer, J. & Nias, A. H. W. *Mutat. Res.* **10**, 67–72 (1970).
22. Maher, V. M., Ouellette, L. M., Mittlestat, M. & McCormick, J. J. *Nature* **258**, 760–763 (1975).
23. Schroy, C. B. & Todd, P. *Mutat. Res.* **33**, 347–355 (1975).
24. Busse, P. M., Bose, S. K., Jones, R. W. & Tolmach, L. J. *Radiat. Res.* **76**, 292–307 (1978).
25. Fogh, J. & Trempe, G. in *Human Tumor Cells In-Vitro* (ed. Fogh, J.) 115–174 (Plenum, New York, 1975).
26. Lawley, P. D. *Prog. Nucleic Acid Res.* **5**, 89–131 (1966).
27. Kohn, K. W., Erickson, L. C., Ewig, R. A. G. & Friedman, C. A. *Biochemistry* **15**, 4629–4637 (1976).
28. Ewig, A. G. & Kohn, K. W. *Cancer Res.* **37**, 2114–2122 (1977).
29. Fornace, A. J. Jr, Little, J. B. & Weischelbaum, R. R. *Biochim. biophys. Acta* **561**, 99–109 (1979).
30. Kohn, K. W., Erickson, L. C., Ewig, R. A. G. & Iqbal, Z. M. *Biochemistry* **13**, 4134–4139 (1974).
31. Makino, F. & Okada, S. *Mutat. Res.* **23**, 387–394 (1974).
32. Tolmach, L. J., Jones, R. W. & Busse, P. M. *Radiat. Res.* **71**, 653–665 (1977).
33. Painter, R. B. & Young, B. R. *Radiat. Res.* **64**, 648–656 (1975).
34. Walters, R. A. & Hildebrand, C. E. *Biochem. biophys. Res. Commun.* **65**, 265–271 (1975).
35. Povirk, L. F. *J. molec. Biol.* **114**, 141–151 (1977).
36. Painter, R. B. *Mutat. Res.* **42**, 299–304 (1977).
37. Lehmann, A. R. *Biophys. J.* **12**, 1316–1325 (1972).
38. Ehmann, U. K., Gehring, U. & Tomkins, G. M. *Biochim. biophys. Acta* **447**, 133–138 (1976).
39. Blumenthal, A. B., Kriegstein, H. J. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **38**, 205–223 (1973).
40. Konze-Thomas, B., Levinson, J. W., Maher, M. & McCormick, J. J. *Biophys. J.* **28**, 315–325 (1979).
41. Fraval, H. N. A. & Roberts, J. J. *Cancer Res.* **39**, 1793–1797 (1979).
42. Hagan, M. P. & Elkind, M. M. *Biophys. J.* **27**, 75–85 (1979).

A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis

Roger J. Watson* & J. Barklie Clements

Institute of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, UK

Controlled transcription of animal virus DNAs provides potentially useful models for elucidating the mechanisms which regulate eukaryotic gene expression. The progressive transcription of the herpes simplex virus type 1 (HSV-1) genome has been described previously^{1–3}. Infection of permissive cells with HSV-1 in the presence of the protein synthesis inhibitor cycloheximide resulted in transcription of a restricted set of virus RNAs, referred to as the immediate early RNAs, which map within certain regions of the virus genome only^{3–5}. Removal of cycloheximide led to the transcription of additional virus DNA sequences, which were expressed during the normal replicative cycle both at early and late times post-infection (before and after the onset of virus DNA replication, respectively) and which map throughout the virus genome³. Previously, we have described a temperature-sensitive mutant of HSV-1 Glasgow strain 17, *ts K*, which accumulated only the immediate early RNAs at the non-permissive temperature (NPT)⁵. Transfer of *ts K*-infected cells from NPT to the permissive temperature (PT), even in the absence of *de novo* protein synthesis, resulted in transcription of the DNA sequences expressed early and late post-infection⁵. This indicated the persistence at NPT of a non-functional immediate early polypeptide which on transfer to PT regained its function, required for progression from the immediate early to early stage of transcription. Here we demonstrate, by analysing the RNAs made in *ts K*-infected cells after transfer from PT to the NPT, that this polypeptide's function is required continuously for synthesis of HSV-1 early and late RNAs, thus identifying a control function essential for the expression of early and late HSV genetic information in eukaryotic cells.

BHK21 C13 cells were infected with *ts K* and maintained at the PT (31 °C) for either 3 or 7 h to allow accumulation of virus specified products. These incubation periods represent early and

late times post-infection, as virus DNA replication initiates at approximately 4 h post-infection in these conditions (data not shown). The cells were then either maintained at the PT or transferred to the NPT (38.5 °C) and, after allowing 30 min for temperature equilibration, were labelled for 3-h periods using ³²P-orthophosphate. RNA isolated from these cells was hybridized to nitrocellulose blot strips containing the separated fragments of HSV-1 DNA generated by endonuclease *Bam*HI digestion, and hybridization was detected by autoradiography. The resultant hybridization patterns (Fig. 1, tracks 2–5) were compared with that of immediate early RNA isolated from *ts K*-infected cells maintained throughout at NPT (Fig. 1, track 6).

RNA samples isolated from *ts K*-infected cells labelled at the PT hybridized to virtually all the *Bam*HI DNA fragments (Fig. 1, tracks 2, 4), and these hybridization patterns were qualitatively identical to those of early and late RNA samples previously observed^{3,5}. In contrast, the hybridization patterns of RNA samples labelled after transfer from the PT to the NPT, both at 3 and 7 h post-infection, were restricted. For example, there was little or no detectable hybridization of these RNA samples to *Bam*HI fragments a, c, d, f, g, h, i, m, p, r, t, uv, w, a' and b' (Fig. 1, tracks 3, 5). These restricted hybridization patterns were indistinguishable from that of *ts K*-infected cells maintained continuously at NPT (Fig. 1, track 6). The DNA regions represented in the RNA synthesized after transfer of *ts K* from the PT to the NPT (Fig. 2) correspond to the major immediate early mRNAs^{6,7}. This suggests that the low abundance of immediate early mRNAs present in cycloheximide-treated cells may represent early transcripts which result from residual protein synthesis in the presence of the drug.

We conclude from the above observations that transfer of *ts K*-infected cells to the NPT resulted in the inactivation of a virus-specified activity required for early and late virus RNA synthesis. To strengthen this conclusion, further control experiments were carried out, first to ensure that the products necessary for early RNA synthesis were present at the time of the early transfer, and second to rule out the possibility that transfer inactivates a host rather than a virus function. Note that neither caveat is probable, as early RNA synthesis was already observed 1 h post-infection at 37 °C (ref. 3) and the wild-type HSV-1 control was found to be transcribed normally at 38.5 °C (ref. 8).

Cells infected with *ts K* were maintained at the PT for 3 h. Fresh medium containing 200 µg ml⁻¹ cycloheximide was then added to the monolayer and, after incubating at the PT for 30 min, cells were labelled in this medium with ³²P-orthophosphate. This treatment thus precluded the accumulation of virus proteins following 3 h post-infection, the time of the early transfer described above.

In a further series of experiments, cells were infected at PT with a *ts*⁺ revertant of *ts K* (*ts K* R1), and cells were labelled at 3.5 h post-infection at either the PT or NPT. RNA samples extracted from these cells were hybridized to the *Bam*HI fragments of HSV-1 DNA, as before. The hybridization patterns obtained with these control RNA samples (Fig. 1, tracks 7–9) were in each instance qualitatively identical to those of *ts K*-infected cells maintained at the PT (Fig. 1, tracks 2, 4).

These results suggested, therefore, that the effect observed on transfer of *ts K*-infected cells from PT to the NPT was the result of cessation of early and late RNA synthesis, and resumption of immediate early RNA synthesis was due to loss of activity of a temperature-sensitive *ts K* product.

The way in which the virus transcriptional function described here acts to effect early and late RNA synthesis is unclear. No modifications of cellular RNA polymerases following HSV-1 infection have been detected by DEAE-Sephadex chromatography, nor were new polymerase activities observed⁹. It has been reported that the activity of a host cell α -amanitin-sensitive RNA polymerase is required for synthesis of all HSV-1 RNA temporal classes¹⁰. Thus, it seems likely that an immediate early virus-coded polypeptide interacts directly with this host cell RNA polymerase, or with the virus DNA template, to enable

* Present address: Laboratory of Molecular Virology, National Cancer Institute, Bethesda, Maryland 20205.

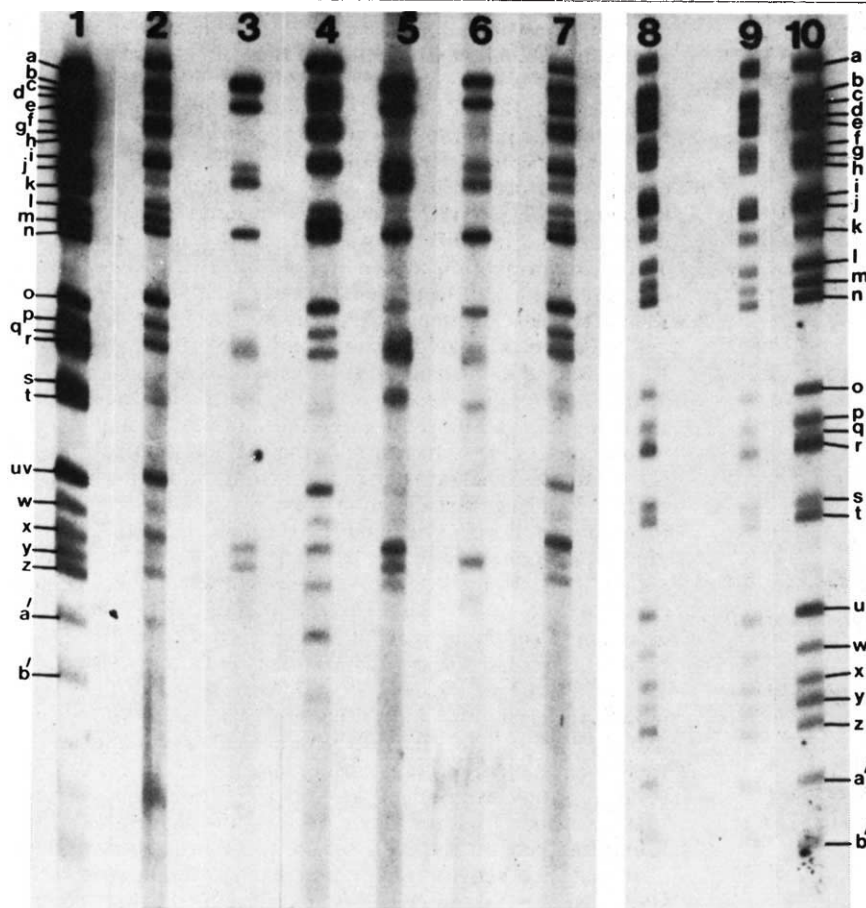


Fig. 1 Autoradiographs of ^{32}P -labelled RNA samples hybridised to the *Bam*HI fragments of HSV-1 DNA. Samples (15 μg) of HSV-1 (Glasgow strain 17) DNA were digested to completion with endonuclease *Bam*HI (Boehringer Mannheim), and the resultant fragments were separated by electrophoresis on a single-well 1% agarose gel⁵. The separated DNA fragments were denatured *in situ* and transferred to nitrocellulose sheets using the Southern blotting technique¹⁴. Strips cut from these blots were incubated with the ^{32}P -labelled nucleic acid samples described below, using hybridisation procedures published previously³. HSV-1 DNA, labelled *in vitro* by nick-translocation⁵, was hybridised to the representative blot strips to indicate the positions of the DNA fragments. Track 1 is the DNA control strip for tracks 2-7, and track 10 the DNA control strip for tracks 8-9. RNA samples were labelled in confluent BHK21 C13 cell monolayers infected at a multiplicity of 25 plaque-forming units per cell with *ts* K (isolated from wild-type Glasgow strain 17)¹⁵ or its revertant *ts* K R1. Cells were incubated for 3 or 7 h at 31 or 38.5 °C, then either transferred from 31 to 38.5 °C or left at the temperature of infection. After allowing 30 min for temperature equilibration, pre-warmed medium containing 0.5 mCi ml⁻¹ ^{32}P -orthophosphate was added, and cells were labelled for 3 h at the appropriate temperature. Total cell RNA was then isolated by proteinase K digestion and phenol extraction, and was purified by isopycnic banding in caesium sulphate gradients⁵. Cells infected with *ts* K (tracks 2-7) were pre-incubated (a) and labelled (b) in the following conditions: track 2, a, 31 °C for 3 h, b, 31 °C; track 3, a, 31 °C for 3 h, b, 38.5 °C; track 4, a, 31 °C for 7 h, b, 31 °C; track 5, a, 31 °C for 7 h, b, 38.5 °C; track 6, 38.5 °C for 3 h, b, 38.5 °C; track 7, a, 31 °C for 3 h, b, 31 °C in medium containing 200 μg ml⁻¹ cycloheximide. Cells infected with *ts* K R1 (tracks 8, 9) were pre-incubated at 31 °C for 3 h, labelled at 31 °C (track 8); pre-incubated at 31 °C for 3 h, labelled at 38.5 °C (track 9).

transcription of early and late virus genes to occur. Such interaction must be independent of virus DNA replication, as the virus-coded transcription-modifying function is required both before and after the onset of virus DNA replication. It could be postulated that the virus function described here is required for stabilization of early and late RNAs rather than for their transcription. This explanation would require that additional transcripts are turned over extremely rapidly at NPT in *ts* K-infected cells, as early and late synthesis was not observed in the 3-h labelling periods reported here.

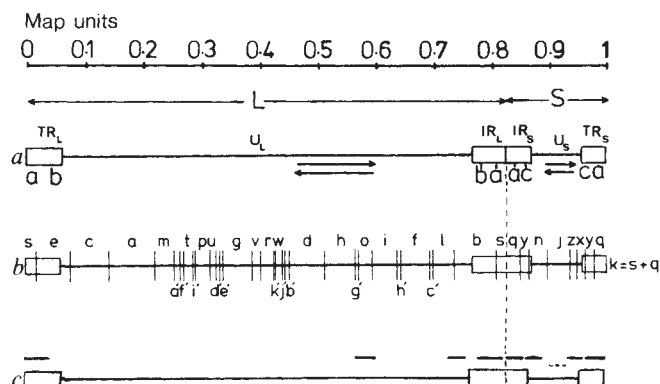


Fig. 2 a, HSV-1 genome organization. Two unique DNA regions U_L and U_S are each flanked by inverted repetitive regions (TR_L/IR_L ; TR_S/IR_S) and joined at IR_L-IR_S . Virion DNA consists of four isomers arising from inversions of U_L and U_S (refs 16, 17) and these inversions are indicated by the directions of arrows. The sequence arrangement for a single strand is indicated, and here the non-inverted terminal redundancy is designated as a. b, Physical map for the HSV-1 DNA fragments generated by the *Bam*HI restriction endonuclease. c, Genome regions represented by *ts* K RNAs labelled after transfer from PT to NPT. Regions represented by abundant RNAs are indicated by solid lines. The summary is based on hybridization to the *Bam*HI fragments (Fig. 1, tracks 3, 5) and to virus DNA fragments generated by other restriction endonucleases⁵.

The *ts* mutation of *ts* K has been mapped in the short reiterated (TR_S/IR_S) region of the virus genome¹¹, a region which specifies an immediate early polypeptide of apparent molecular weight 175,000 (Vmw 175)^{6,12}. This polypeptide exhibited abnormal properties at NPT in *ts* K-infected cells, characterized by a failure to be fully processed and by a greater retention in the cytoplasm than in the nucleus¹³. All the abnormal properties were reversed by transfer of cells from the NPT to the PT, suggesting that they have functional significance and indicating the possibility that the temperature-sensitive *ts* K lesion is within Vmw 175.

The HSV-1 mutant *ts* K provides an attractive model for analysing the control of gene expression in infected eukaryotic cells. The development of appropriate *in vitro* transcriptional systems should enable the function required for HSV-1 transcription to be examined more readily.

We thank Professor J. H. Subak-Sharpe for his continued interest in this work, and D. Dargan and J. H. Subak-Sharpe for *ts* K R1. R. J. W. was the recipient of a short-term grant from the MRC.

Received 14 December 1979; accepted 5 March 1980.

1. Frenkel, N. & Roizman, B. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2654-2658 (1972).
2. Swanson, R. I. & Wagner, E. K. *Virology* **60**, 522-533 (1974).
3. Clements, J. B., Watson, R. J. & Wilkie, N. M. *Cell* **12**, 275-285 (1977).
4. Jones, P. C., Howard, G. S. & Roizman, B. *J. Virol.* **21**, 268-276 (1977).
5. Watson, R. J. & Clements, J. B. *Virology* **91**, 364-379 (1978).
6. Watson, R. J., Preston, C. M. & Clements, J. B. *J. Virol.* **31**, 42-52 (1979).
7. Clements, J. B., McLauchlan, J. & McGeoch, D. J. *Nucleic Acids Res.* **7**, 77-91 (1979).
8. Watson, R. J. thesis, Univ. Glasgow (1979).
9. Lowe, P. A. *Virology* **86**, 577-580 (1979).
10. Costanzo, F., Campadelli-Fiume, G., Foa-Tomasi, L. & Cassai, E. *J. Virol.* **21**, 996-1001 (1977).
11. Stow, N. D., Subak-Sharpe, J. H. & Wilkie, N. M. *Virology* **90**, 1-11 (1978).
12. Preston, V. G. *et al. J. Virol.* **28**, 499-517 (1979).
13. Preston, C. M. *J. Virol.* **32**, 357-369 (1979).
14. Southern, E. M. *J. molec. Biol.* **98**, 503-533 (1975).
15. Crombie, I. K. thesis, Univ. Glasgow (1975).
16. Hayward, G. S., Frenkel, N. & Roizman, B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4243-4247 (1975).
17. Delius, H. & Clements, J. B. *J. gen. Virol.* **33**, 125-133 (1976).

Differential stimulation of capped mRNA translation *in vitro* by cap binding protein

N. Sonenberg*, H. Trachsel†, S. Hecht‡ & A. J. Shatkin

Roche Institute of Molecular Biology, Nutley, New Jersey 07110

†Friedrich Miescher-Institut, CH-4002 Basel, Switzerland

‡Departments of Chemistry and Biology, University of Virginia, Charlottesville, Virginia 22901

The cap structure, m⁷GpppN, which is present at the 5' end of most eukaryotic mRNAs, facilitates the initiation of protein synthesis¹. Rabbit reticulocytes and mouse L and ascites cells have been shown to contain a protein (molecular weight (MW) 24,000) that specifically recognizes the cap of reovirus mRNA and several other capped mRNAs². That protein partially copurifies with eIF-3 and eIF-4B (refs 2, 3), two eukaryotic initiation factors involved in mRNA binding to ribosomes⁴. We have isolated essentially homogeneous 24,000-MW polypeptide which stimulated translation of capped mRNAs *in vitro*⁵. Consistent with those findings, we now report that among the many protein synthesis factors tested, preparations of eIF-3 and eIF-4B that also contained the 24,000-MW cap binding protein (CBP) markedly increased translation in HeLa cell extracts of capped mRNAs but not of naturally uncapped viral messengers. Moreover, removal of the 24,000-MW protein diminished the ability of eIF-3 to stimulate translation of capped mRNA. Purified CBP, which had no effect on uncapped satellite tobacco necrosis virus (STNV) RNA, stimulated translation of this RNA after addition of a cap. The protein also relieved translational competition between capped and naturally uncapped mRNAs in favour of the capped species. These findings support the hypothesis that the CBP participates in the regulation of eukaryotic mRNA translation in normal and virus-infected cells⁶.

Cap binding protein purified by affinity chromatography in m⁷GDP-Sepharose⁵ stimulated translation in HeLa cell extracts of Sindbis virus mRNA which is capped⁷ (Fig. 1a) but not encephalomyocarditis (EMC) virus RNA, an uncapped message⁸ (Fig. 1b). The differential effect on Sindbis mRNA as well as on capped reovirus and globin mRNAs was greater at higher concentrations of CBP⁵, promoting increased synthesis of Sindbis capsid and its B₁ precursor protein⁹, reovirus structural proteins¹⁰ and globin as determined by polyacrylamide gel electrophoresis of the radiolabelled products (data not shown). By contrast, there was no effect on EMC virus RNA and another uncapped message¹¹, STNV RNA. This was apparently not due to mRNA saturation because no stimulation was observed when low concentrations of EMC virus RNA were used (Fig. 1d), whereas Sindbis mRNA translation was increased at several concentrations (Fig. 1c).

The effect of CBP was tested in extracts directed by mixtures of capped and uncapped RNAs that presumably were in competition for components necessary for translation. Alfalfa mosaic virus (AMV)-4 RNA is a capped mRNA that codes for viral coat protein¹² (Fig. 2, lane 2). In the presence of STNV RNA, a relatively weak message in extracts of HeLa cells as compared to wheat germ¹³, both STNV products (Fig. 2, lane 3) and the larger AMV coat protein were made (Fig. 2, lane 4). The yield of AMV protein was increased in response to the RNA mixture, possibly due to protection against RNase digestion by the higher total concentration of mRNA. Addition of CBP increased further the translation of AMV RNA whereas STNV RNA translation decreased slightly (Fig. 2, lanes 5–7). Similarly,

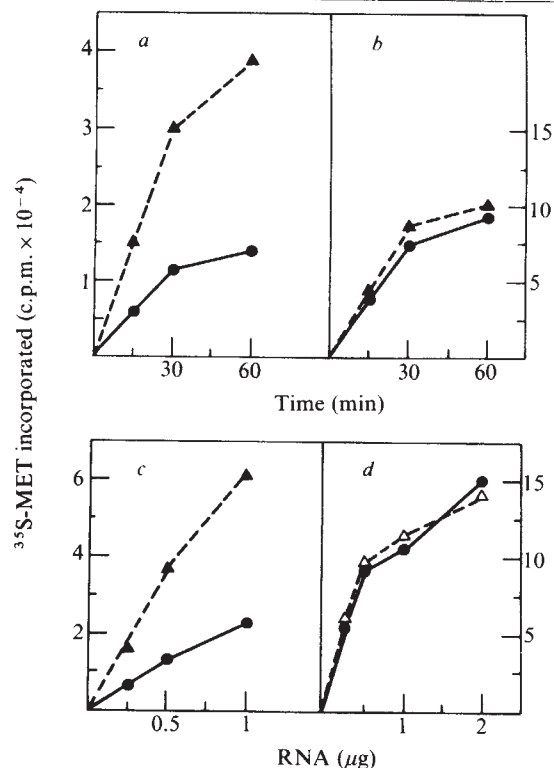


Fig. 1 Effect of CBP on translation of Sindbis virus mRNA and EMC virus RNA in HeLa cell-free extracts. Conditions of *in vitro* protein synthesis were as described previously⁵. ³⁵S-Methionine (16 μCi, specific activity 643 Ci mmol⁻¹) incorporation into acid-precipitable products was assayed in 5-μl aliquots of 25-μl incubation mixtures. Sindbis virus mRNA²⁵, EMC virus RNA²⁶ and rabbit reticulocyte 24,000-MW CBP⁵ were present at ~0.5 μg, 0.5 μg and 30 ng, respectively, or at the indicated amounts. Levels of radioactivity incorporated in the absence of exogenous mRNAs (4,695 c.p.m. without and 4,970 c.p.m. with CBP) were subtracted. Panels a, c: +Sindbis virus mRNA; b, d: +EMC virus RNA. Dotted line: +CBP.

in a mixture of globin and STNV RNAs, the presence of CBP stimulated globin formation and, at the highest level, decreased the synthesis of STNV products (Fig. 2, lanes 8–11). The results suggest that the CBP discriminates against naturally uncapped mRNAs in favour of capped mRNAs in a mixture of competing molecules. Consistent with this suggestion, formation of the spectrum of EMC virus products was diminished and AMV RNA translation was increased by addition of the 24,000-MW protein (Fig. 2, lanes 12–14).

To explore the possibility that the presence of co-purified CBP accounts for some previous reports of mRNA discriminatory activity by eukaryotic initiation factors^{14–18}, rabbit reticulocyte factors were tested for their ability to stimulate Sindbis as compared with EMC RNA. Only eIF-4B (Table 1, expt 1) and eIF-3 (Table 1, expt 2) had discriminatory activity, increasing the translation of Sindbis mRNA 9–12-fold more than EMC virus RNA. Factor eIF-3 contains CBP unless washed with 0.5 M KCl⁶. Consequently, we also tested a preparation of salt-washed eIF-3 which had been shown to be free of the 24,000-MW polypeptide by a cap cross-linking assay². This eIF-3 failed to stimulate translation of Sindbis mRNA (Table 1, expt 1) although it retained eIF-3 activity in a reconstituted mammalian cell-free protein synthesizing system⁶. Another eIF-3 that contained one-third as much CBP as the preparation used in expt 2 (Table 1) gave an intermediate ratio of 4.3 for stimulation of Sindbis compared with EMC virus RNA translation (data not shown). Thus, the presence of 24,000-MW protein is correlated with the ability of a factor preparation to stimulate differentially capped mRNA translation *in vitro*.

* Present address: McIntyre Medical Sciences Building, McGill University, Montreal, Canada.

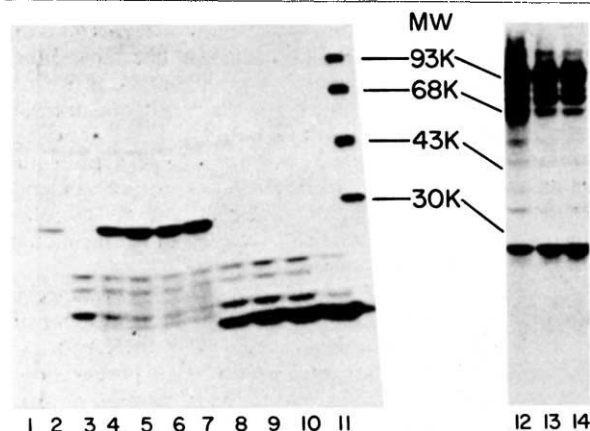


Fig. 2 The effect of CBP on translation of mixtures of capped and uncapped mRNAs by HeLa cell-free extracts. ^{35}S -Methionine-labelled products synthesized during a 60-min incubation were analysed by polyacrylamide gel electrophoresis and fluorography^{27,28}. Reaction mixtures included no added RNA (lane 1) or 2 μg of each of the following RNAs: AMV-4 (lane 2), STNV (lane 3), AMV-4 + STNV and 0, ~15, 30 and 45 ng CBP (lanes 4–7, respectively); globin + STNV and 0, ~15, 30 and 45 ng CBP (lanes 8–11); and, in a separate experiment, AMV-4 + EMC and 0, ~15 and 30 ng CBP (lanes 12–14).

To determine if addition of a 5' cap to a naturally uncapped RNA would result in an effect of CBP, STNV RNA was translated before and after exposure to capping enzymes purified from vaccinia virus^{19,20}. Because STNV RNA is a weak messenger in the HeLa system, much of the incorporated ^{35}S -methionine appeared in a prominent cellular protein (Fig. 3, lane 1). After treatment with vaccinia capping enzymes, the STNV RNA which included capped (~20%) and uncapped (~80%) molecules was less effectively translated (Fig. 3, lane 2) than untreated RNA (Fig. 3, lane 4). This was probably due to nicking of the RNA during the capping reaction, as most of the products directed by treated STNV RNA migrated slightly faster than viral capsid protein (MW ~22,000)²¹. The smaller polypeptide (MW ~18,000) which was also synthesized in response to uncapped STNV RNA but in lower amounts than capsid (Fig. 3, lane 4) presumably represents prematurely terminated capsid protein²¹. When CBP was added to an incubation mixture directed by the STNV RNA that contained

some capped molecules, it increased the synthesis of the viral polypeptide products (compare Fig. 3, lanes 2 and 3). By contrast, no stimulation occurred with untreated STNV RNA (compare Fig. 3, lanes 4 and 5).

Many investigators have reported that certain eukaryotic initiation factors can relieve translational competition between two different mRNAs in mammalian cell-free protein synthesising systems. For example, Golini *et al.*¹⁴ reported that eIF-4B can overcome the competitive advantage of EMC virus RNA over rabbit globin mRNA and other capped cellular mRNAs. Nudel *et al.*²² described a protein with a similar property, that is, the ability to relieve translational competition between the RNA of mengovirus, another picornavirus, and globin mRNA (and between α - and β -globin capped mRNAs). Kabat and Chappell¹⁵ also observed that eIF-4B could overcome the preferential translation of rabbit β -globin mRNA relative to α -globin mRNA. We found that purified CBP preferentially stimulated translation of several capped mRNAs at the expense of the uncapped RNAs of EMC virus and STNV. Because we had previously observed that initiation factors eIF-4B and eIF-3 contain variable amounts of the 24,000-MW protein², we also tested these two initiation factors and found that both stimulated Sindbis virus capped mRNA severalfold but were without significant effect on EMC virus uncapped RNA. The differential stimulatory activity of eIF-3 for capped mRNA translation was correlated with the amount of 24,000-MW CBP in the different factor preparations. These observations agree with the report²³ that cell-free extracts of mouse L or ascites cells translate well the uncapped RNA of mengovirus, but require supplementation by reticulocyte factors for efficient translation of globin mRNA. We obtained similar results with ascites cell extracts directed by reovirus mRNA and EMC virus

Table 1 Effect of protein synthesis factors on translation of capped and uncapped RNA

Factor added	Incorporation (c.p.m.)		Fold-stimulation		Ratio of stimulation	
					Sindbis RNA	
	Sindbis RNA	EMC RNA	Sindbis RNA	EMC RNA	EMC RNA	EMC RNA
Expt 1*						
None	1,754	8,830	—	—	—	—
eIF-2 (3.5 μg)	3,952	9,527	2.2	1.1	2.0	—
eIF-3	1,218	12,757	0.7	1.4	0.5	—
(5 μg of 0.5 M KCl washed)						
eIF-4A (6 μg)	9,851	19,550	5.6	2.2	2.5	—
eIF-4B (1.5 μg)	23,919	12,943	13.6	1.5	9.1	—
eIF-4C + D (2 μg)	1,353	4,021	0.8	0.5	1.6	—
eIF-5 (0.5 μg)	1,105	14,411	0.6	1.6	0.4	—
EF-2 (2 μg)	1,699	15,910	1.0	1.8	0.6	—
Expt 2						
None	7,880	87,652	—	—	—	—
EF-1 (1 μg)	9,814	81,397	1.2	0.9	1.3	—
eIF-3	96,083	79,598	12.2	0.9	12.3	—
(4 μg of 0.1 M KCl washed)						

RNAs (0.5 μg 25 μl^{-1}) were incubated in HeLa cell-free extracts for 60 min with ^{35}S -methionine (24 μCi , specific activity 893 Ci mmol^{-1}) as radioactive precursor. Acid-precipitable radioactivity was assayed as in Fig. 1 (refs 5, 6). Factors were estimated to be 80–90% pure by polyacrylamide gel electrophoresis^{3,6}, and the amounts used were above saturating levels for the reticulocyte reconstituted translation system⁴.

* For expts 1 and 2 the subtracted levels of endogenous synthesis were 3,724 and 8,877 c.p.m., respectively.

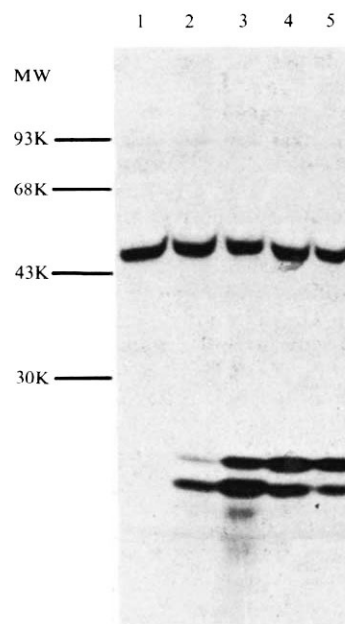


Fig. 3 Effect of CBP on the translation of capped and uncapped STNV RNA. STNV RNA was incubated for 20 min at 37 °C as described previously^{19,20} in the presence of GTP, ^3H -methyl-S-adenosylmethionine and guanylyl- and methyl-transferase activities solubilized and partially purified from vaccinia virus. Analysis of a P_1 nuclease digest of the ^3H -labelled RNA by high voltage electrophoresis indicated that 20% of the 5' termini had been converted to the m⁷G-capped form. ^{35}S -Methionine-labelled polypeptide products directed by the RNAs were analyzed by polyacrylamide gel electrophoresis. Lane 1: no added RNA; 2: 2 μg 'capped' STNV RNA; 3: as lane 2 plus ~45 ng CBP; 4: 2 μg STNV RNA (not capped); 5: as lane 4 plus ~45 ng CBP. The yields of acid-precipitable radioactivity in c.p.m. per 5- μl aliquot of incubation mixture for lanes 1–5 were respectively, 5,528, 7,592, 11,631, 11,250 and 10,951; 20- μl aliquots were applied to the gel.

RNA (unpublished results). Furthermore, the capacity of eIF-4B to restore capped mRNA translation in poliovirus-infected HeLa cell extracts is probably due to co-purification of 24,000-MW CBP with the 80,000-MW main component in eIF-4B preparations⁶. Poliovirus apparently exerts its shut-off effect on host cell protein synthesis by inactivation of the 24,000-MW protein^{6,17}. By this mechanism the uncapped viral RNA, which is presumably independent of or less dependent on CBP, redirects the translational machinery of infected cells. The *in vitro* effect of CBP on capped mRNA translation may reflect a key regulatory role in poliovirus-infected cells.

Shut-off of host protein synthesis by functional inactivation of CBP might not be restricted to picornaviruses that contain uncapped RNA but may be a more general eukaryotic control mechanism. For example, it was recently reported²⁴ that mouse L cell extracts prepared at late times after reovirus infection

translate uncapped, but not capped viral mRNA. This is in sharp contrast to uninfected cell extracts which, like other eukaryotic protein synthesizing systems¹, utilize capped mRNA preferentially. Thus, reovirus infection may modify the translational machinery of L cells in a way that results in a switch from translation of capped to uncapped mRNA, perhaps by inactivation of CBP with parallel development of cap-independent translation mechanisms analogous to events in poliovirus-infected HeLa cells. Other viruses may re-direct the metabolism of infected cells in similar ways.

We thank A. LaFiandra and M. Morgan for assistance, Drs G. Kramer and W. Merrick for preparations of protein synthesis factors, Drs J. Clark and A. Marcus for STNV RNA and AMV-4 RNA, and Dr K. Rupprecht for help with the preparation of the affinity resin. Part of this work was supported by grant MA-7214 from the Canada MRC to N.S.

Received 17 December 1979; accepted 29 February 1980.

- Shatkin, A. J. *Cell* **9**, 645-653 (1976).
- Sonenberg, N., Morgan, M. A., Merrick, W. C. & Shatkin, A. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4843-4847 (1978).
- Bergmann, J. E., Trachsel, H., Sonenberg, N., Shatkin, A. J. & Lodish, H. F. *J. biol. Chem.* **254**, 1440-1443 (1979).
- Trachsel, H., Erni, B., Schreier, M. H. & Staehelin, T. *J. molec. Biol.* **116**, 755-767 (1977).
- Sonenberg, N., Rupprecht, K. M., Hecht, S. M. & Shatkin, A. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4345-4349 (1979).
- Trachsel, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 770-774 (1980).
- Dubin, D. T., Timko, K., Gillies, S. & Stollar, V. *Virology* **98**, 131-141 (1979).
- Frisby, D., Eaton, M. & Fellner, P. *Nucleic Acids Res.* **3**, 2771-2787 (1976).
- Simmons, D. T. & Strauss, J. H. *J. molec. Biol.* **86**, 397-409 (1974).
- Shatkin, A. J. & Both, G. W. *Cell* **7**, 305-313 (1976).
- Wimmer, E., Chang, A. Y., Clark, J. M. Jr & Reichman, M. E. *J. molec. Biol.* **38**, 59-73 (1968).
- Van Vloten-Doting, L. & Jaspars, E. M. J. *Comprehensive Virol.* **11**, 1-53 (1977).
- Herson, D., Schmidt, A., Seal, S., Marcus, A. & Van Vloten-Doting, L. *J. biol. Chem.* **254**, 8245-8249 (1979).
- Golini, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **73**, 3040-3044 (1976).
- Kabat, D. & Chappell, J. R. *J. biol. Chem.* **252**, 2684-2690 (1977).
- Shafritz, D. A. *et al. Nature* **261**, 291-294 (1976).
- Rose, J. K., Trachsel, H., Leong, K. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2732-2736 (1978).
- Gette, W. R. & Heywood, S. M. *J. biol. Chem.* **254**, 9879-9885 (1979).
- Moss, B. *Biochem. biophys. Res. Commun.* **74**, 374-383 (1977).
- Sonenberg, N., Shatkin, A. J., Ricciardi, R. P., Rubin, M. & Goodman, R. M. *Nucleic Acids Res.* **5**, 2501-2512 (1978).
- Salvato, M. S. & Fraenkel-Conrat, H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2288-2292 (1977).
- Nudel, U., Lebleu, B. & Revel, M. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2139-2144 (1973).
- Lebleu, B., Nudel, U., Falcoff, E., Prives, C. & Revel, M. *FEBS Lett.* **25**, 97-103 (1972).
- Skup, D. & Millward, S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 152-156 (1980).
- Cancedda, R. & Shatkin, A. J. *Eur. J. Biochem.* **94**, 41-50 (1979).
- Eggen, K. L. & Shatkin, A. J. *J. Virol.* **9**, 636-645 (1972).
- Laemmli, U. K. & Favre, J. *J. molec. Biol.* **80**, 575-599 (1973).
- Sonenberg, N. & Shatkin, A. J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4288-4292 (1977).

Construction *in vitro* and rescue of a thymidine kinase-deficient deletion mutation of herpes simplex virus

James R. Smiley

Pathology Department, McMaster University Medical Center, 1200 Main St W., Hamilton, Ontario, Canada L8N 3Z5

The ability to introduce defined mutations at pre-selected sites on DNA molecules *in vitro* allows one to test hypotheses about the function of specific DNA sequences. The genome of herpes simplex virus type 1 (HSV-1), consisting of 150 kilobase pairs, is probably too large to allow the direct introduction of mutations *in vitro*. Therefore, to apply these techniques to the study of HSV genes, the feasibility of a two-step procedure was examined. A defined mutation was first introduced into a small segment of HSV-1 DNA which had previously been cloned in the *Escherichia coli* plasmid pBR322; the mutation was then incorporated into the genome of infectious HSV-1 by DNA-mediated marker rescue. The approach was tested using the viral gene for thymidine kinase (TK)¹⁻³. TK was chosen for several reasons. First, TK is dispensable for viral infectivity in most conditions of cell culture, and efficient selection systems are available for both the TK⁻ and TK⁺ phenotypes^{1,3}. This facilitates the recovery of TK-deficient mutants and allows rapid and detailed genetic analysis^{3,4}. Second, Stow *et al.*⁵ have shown that TK⁻ DNA sequences are readily rescued from DNA fragments into infectious virus. Third, purified HSV-1 TK DNA is able to convert TK-deficient mammalian cells to a TK⁺ phenotype⁶, so that the expression of the viral gene can be studied in both the presence and absence of other regulatory viral gene products. Thus, the TK gene provides a favourable system for an examination of the role of specific DNA sequences in the expression of genes in eukaryotic cells. The results presented here demonstrate that an 800-base pair deletion mutation at the TK locus, generated *in vitro*, was rescued at a high frequency, producing a TK-deficient deletion mutant of HSV-1.

A 3.5-kilobase *Bam*HI fragment of HSV-1 DNA, containing the TK gene⁶, has been inserted at the *Bam*HI site of pBR322 (ref. 7). The resulting recombinant plasmid, pX1 (Fig. 1), was used to isolate variant plasmids bearing defined deletions which potentially inactivate the viral gene. A deletion mapping entirely within the TK coding sequence was sought, so as to minimize the probability of also inactivating any neighbouring essential HSV genes. Approximately 1.2 kilobases of DNA are required to specify the 44,000-molecular weight TK polypeptide³. As the TK-transforming activity of the 3.5-kilobase *Bam*HI fragment resides the largest *Pvu*II subfragment⁸, and is abolished by cleaving the DNA with *Eco*RI⁶, it seemed likely that a deletion of the 800-base pair *Pst*I fragment would map entirely within TK. Accordingly, a plasmid (pd2, Fig. 1) bearing the deletion was constructed (Fig. 1 legend).

To determine whether the deletion borne by pd2 could be recovered as a viable TK-deficient mutation of HSV-1, Vero cells were transfected with a co-precipitate formed with wild-type HSV-1 43 TK⁺ (ref. 7) membrane-stripped nucleocapsids⁹, 7.5 µg ml⁻¹ purified circular pd2 DNA and 10 µg ml⁻¹ carrier cell DNA, using the calcium phosphate technique of Graham and Van der Eb¹⁰. The virus recovered from the infected cultures was then scored for the fraction of TK-deficient mutants by plaque titration on Vero cells in the presence and absence of 6 µg ml⁻¹ 5-bromodeoxyuridine. The presence of pd2 DNA in transfections led to a fivefold increase in the frequency of recovery of TK-deficient virus (3×10^{-3} compared with 6×10^{-4} in transfections with nucleocapsids and carrier DNA alone). Seven TK-deficient plaques selected from the progeny of the mixed infection by plaque purification in the presence of 5-bromodeoxyuridine were amplified in Vero cells, then screened for deletions by examining the *Eco*RI cleavage pattern of ³²P-labelled viral DNA. Five of the seven isolates bore the expected 800-base pair deletion within the *Eco*RI N fragment¹¹ of viral DNA. One such isolate, HSV-1 TK d2, was plaque purified again, expanded and examined in greater detail.

Figure 2 compares the *Bam*HI and *Eco*RI cleavage patterns of the DNAs of HSV-1 43 TK⁺, HSV-1 TK d2 and the corresponding recombinant plasmids pX1 and pd2. HSV-1 TK d2 DNA differs from that of its TK⁺ parent in that the normal

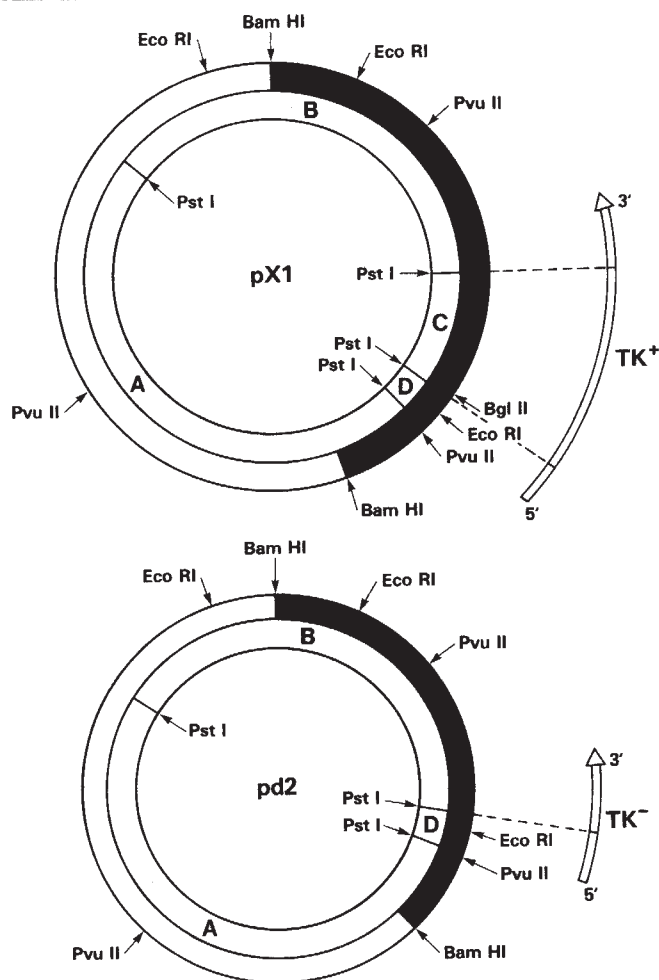


Fig. 1 Construction of pd2 from pX1. The deletion-bearing plasmid pd2 was derived from a partial *Pst*I digest of pX1 DNA. pX1 DNA (4 μ g) was incubated with 0.6 units of *Pst*I (New England Biolabs) in 20 μ l containing 50 mM KCl, 10 mM MgCl₂ and 10 mM Tris-HCl, pH 7.5, for 15 min at 37 °C, conditions which maximized the yield of doubly cleaved molecules. The reaction was terminated by the addition of EDTA to 20 mM, and the DNA was purified by extraction first with phenol then with ether, and concentrated by precipitation with alcohol. The partially digested DNA (100 ng) was incubated at 4 °C for 16 h in 20 μ l containing 66 mM Tris-HCl, pH 7.6, 6.6 mM MgCl₂, 10 mM dithiothreitol, 0.4 mM ATP and 0.25 units of T4 DNA ligase (New England Biolabs). Following ligation, 5 units of *Bgl*II were added to the reaction mixture to inactivate any molecules retaining the *Pst*IC fragment. After a further incubation for 2 h at 37 °C, the DNA was used directly to transform CaCl₂-treated *E. coli* LE392 (ref. 7). Ampicillin-resistant transformants were obtained on LB plates containing 20 μ g ml⁻¹ ampicillin. The plasmid DNA from six colonies was analysed for deletions by cleavage with *Pst*I and *Eco*RI. One plasmid (pd2), lacking only the *Pst*IC fragment, was identified. The structure of pd2 DNA was further verified by examining the fragments produced by cleavage with *Bam*HI, *Sma*I and double cleavage with *Eco*RI plus *Pvu*II. The HSV-1 DNA inserts are indicated by the filled portions of the outer circle. The locations of restriction endonuclease cleavage sites on pX1 DNA are from Enquist *et al.*⁷, the approximate map location of the TK gene was inferred from the data of Colbere-Garapin *et al.*⁸ and Wigler *et al.*⁶, and the polarity of transcription of TK is from Smiley *et al.*⁴.

3.5-kilobase *Bam*HI TK fragment is replaced by the deletion-bearing 2.7-kilobase fragment derived from pd2. This replacement of wild-type sequences by mutant DNA is partially obscured by the coincidental fixation in HSV-1 TK d2 of a smaller version of the *Bam*HI P fragment of HSV DNA, which now co-migrates with the wild-type TK fragment. The *Bam*HI P fragment, which in HSV-1 43 TK⁺ DNA runs as a broad band

just above the TK fragment, is derived from the end of the short segment of HSV DNA¹², a region of the genome known to acquire deletion and insertion mutations at a high frequency⁹. The length of this fragment varies both between HSV-1 strains¹² and between different plaque-purified stocks of the same strain (M. J. Wagner and W. C. Summers, personal communication). The replacement of wild-type TK DNA sequences by mutant DNA is more directly demonstrated by the *Eco*RI cleavage patterns: the wild-type 2.3-kilobase *Eco*RI N fragment¹¹ of viral DNA is replaced by the mutant 1.5-kilobase fragment derived from pd2 DNA. These data demonstrate that the deletion carried by pd2, generated *in vitro*, has been incorporated into the HSV-1 TK d2 genome. This deletion abolishes the ability of the mutant virus to induce HSV-1 TK activity following infection of TK-deficient LtA cells¹³ (Fig. 3).

As neither pBR322 nor wild-type TK DNA sequences are found in HSV-1 TK d2 DNA, it follows that the mutant virus arose by two cross-over events between viral and plasmid DNA. It is not clear whether both cross-overs occurred during the

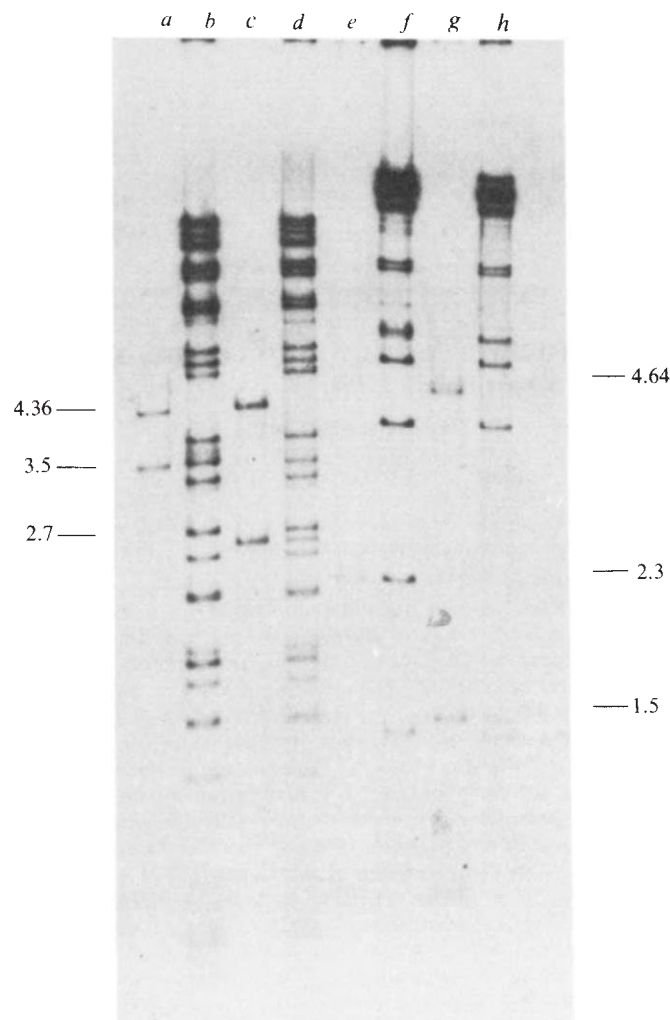


Fig. 2 Restriction endonuclease cleavage pattern of HSV-1 d2 DNA. ³²P-labelled viral DNA was extracted from viral nucleocapsids as previously described⁹. Plasmid DNA was labelled *in vitro* by nick translation using [α -³²P]dCTP (NEN) and DNA polymerase I (Boehringer-Mannheim) according to Maniatis *et al.*¹⁴. No DNase I was added. The DNA was then cleaved with *Eco*RI (Boehringer-Mannheim) or *Bam*HI in a reaction volume of 10 μ l, containing 50 mM KCl, 10 mM MgCl₂ and 10 mM Tris-HCl, pH 7.5. DNA fragments were resolved by electrophoresis through a 15 $\times$ 30-cm vertical 1.0% agarose gel as previously described⁹. *Bam*HI digests: a, pX1; b, HSV-1 43 TK⁺; c, pd2; d, HSV-1 TK d2. *Eco*RI digests: e, pX1; f, HSV-1 43 TK⁺; g, pd2; h, HSV-1 TK d2.

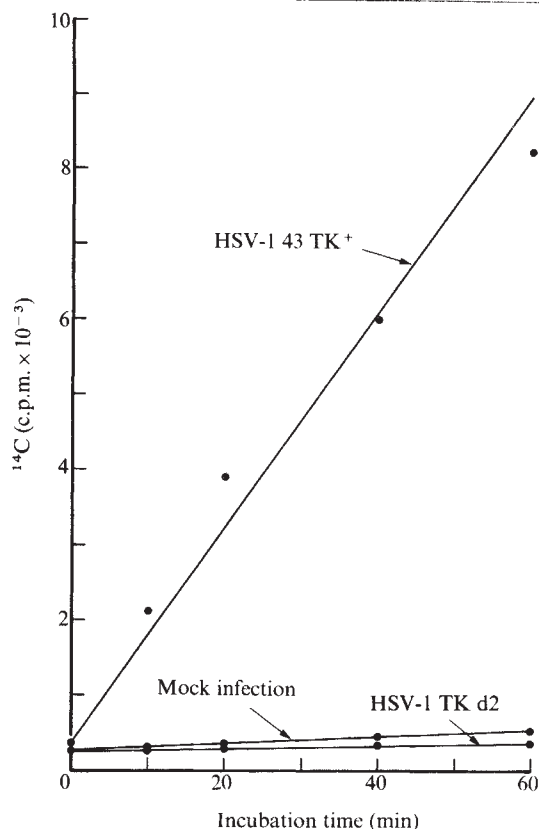


Fig. 3 TK activity induced by HSV-1 TK⁺ and TK d2. 2×10^7 LtA cells were infected with 10^8 plaque-forming units of the indicated virus and collected 18 h post-infection. Extracts were assayed for TK activity using ^{14}C -thymidine, exactly as previously described³.

initial interaction between the plasmid and viral DNA, or whether an initial single cross-over event between the circular plasmid and linear viral genomes gave rise to an unstable TK⁺/TK⁻ heterozygote which subsequently segregated TK⁻ homozygotes by unequal crossing over between the resulting direct repeats of viral DNA. Because TK⁺/TK⁻ heterozygotes would score as TK⁺, they would not have been recovered in these experiments, in which the selection was for TK-deficient virus. Experiments in which the TK⁺ phenotype is rescued from pX1 by HSV-1 TK d2 may allow the recovery of such heterozygotes. These experiments are in progress.

The region of HSV-1 DNA encompassed by the d2 deletion is now open to exhaustive analysis by site-directed mutagenesis, as any alteration of the DNA sequences mapping within the deletion can be examined for consequent effects on the TK gene. If such an alteration inactivates TK, it can be rescued as a TK-deficient mutation; otherwise, it can be used to convert HSV-1 TK d2 to TK⁺. The aim of this research is to extend this analysis to the DNA sequences which encode signals governing the synthesis and processing of TK mRNA. In addition, this and other deletion mutations at the TK locus of HSV-1 will be of value for correlating the physical and genetic maps of the TK gene, and for studies of the control of TK expression during lytic infection and in biochemically transformed cells. *In vitro* mutagenesis using cloned segments of HSV DNA will also be useful in the study of other viral genes.

I thank H. Swan for technical assistance, S. Bacchetti, W. E. Rawls and R. J. Redfield for critical readings of the manuscript, and W. C. Summers for *Bam*HI. This research was supported by the National Cancer Institute of Canada. Experiments with recombinant DNA were carried out in accordance with the MRC of Canada guidelines. The author is a Research Scholar of the NCI of Canada.

Received 1 February; accepted 19 March 1980.

1. Dubbs, D. R. & Kit, S. *Virology* **22**, 493–502 (1964).
2. Klemperer, H. G., Haynes, G. R. & Watson, D. H. *Virology* **31**, 120–128 (1967).
3. Summers, W. P., Wagner, M. & Summers, W. C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4081–4084 (1975).
4. Smiley, J. R., Wagner, M. J., Summers, W. P. & Summers, W. C. *Virology* (in the press).
5. Stow, N. D., Subak-Sharpe, J. H. & Wilkie, N. M. *J. Virol.* **28**, 182–192 (1978).
6. Wigler, M. *et al. Cell* **11**, 223–232 (1977).
7. Enquist, L. W., Van de Woude, G. F., Wagner, M. J., Smiley, J. R. & Summers, W. C. *Gene* **7**, 335–342 (1979).
8. Colbere-Garapin, F., Chousterman, S., Horodniceau, F., Kourilsky, P. & Garapin, A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3755–3759 (1979).
9. Wagner, M. J. & Summers, W. C. *J. Virol.* **27**, 374–387 (1978).
10. Graham, F. L. & Van der Eb, A. J. *Virology* **52**, 456–467 (1973).
11. Skare, J. & Summers, W. C. *Virology* **76**, 581–599 (1977).
12. Locker, H. & Frenkel, N. J. *Virol.* **32**, 429–441 (1979).
13. Wigler, M., Pellicer, A., Silverstein, S. & Axel, R. *Cell* **14**, 725–731 (1978).
14. Maniatis, T., Jeffrey, A. & Kleid, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).

Heterogeneity in the molecular basis of hereditary persistence of fetal haemoglobin

Dorothy Tuan, Mary J. Murnane, J. K. deRiel & Bernard G. Forget

Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut 06510

In hereditary persistence of fetal haemoglobin (HPFH), there is a uniformly high level of synthesis of fetal haemoglobin (Hb F: $\alpha_2\gamma_2$) in all red cells of affected adults. Other hereditary disorders of human haemoglobin synthesis may also be associated with persistent synthesis of Hb F into adult life, but usually at a lower level and in only a selected population of red cells. Previous studies in the common type of HPFH that occurs in blacks have demonstrated the presence of a total deletion of the linked β - and δ -globin genes of the adult haemoglobins Hb A ($\alpha_2\beta_2$) and Hb A₂ ($\alpha_2\delta_2$)^{1–3}, including approximately 4 kilobases of the DNA flanking the 5' end of the δ gene. In contrast, $\delta\beta$ -thalassaemia, in which Hb F synthesis is not as elevated or uniform as in HPFH, is associated with a partial deletion of the δ gene that leaves intact the 5'-flanking DNA of the gene^{1–5}. These results provided support for previous theories⁶ that regulatory sequences involved in the suppression of γ -globin gene expression might be located between the γ - and δ -globin genes, and that deletion of these sequences might be responsible for the continued expression of γ -globin genes in HPFH. To extend these observations, gene mapping studies of the non- α -globin genes have now been carried out in two additional individuals with HPFH using restriction endonuclease digestion and gel blotting of total cellular DNA according to the method of Southern¹⁶. The results indicate that the molecular defects associated with HPFH are heterogeneous even in cases with similar phenotypes such as in the common variety of HPFH that occurs in blacks.

In addition to the previously studied^{1,2} case of HPFH in an American black (black HPFH-1), we analysed the DNA from two other individuals with HPFH: a case of homozygous HPFH from Ghana (black HPFH-2) with the common type of black HPFH in which both $^G\gamma$ - and $^A\gamma$ -globin chains are synthesized, and a heterozygote for the Greek type of HPFH in which $^A\gamma$ -globin chain synthesis vastly predominates⁷. The results of our restriction endonuclease mapping of these DNAs are summarized in Fig. 1.

In normal DNA, restriction endonuclease *Bcl*I cleaves outside the two γ -globin genes and generates a single 20-kilobase fragment that hybridizes to the γ -gene probe⁸. DNA from all three cases of HPFH, when cleaved by *Bcl*I, yields a single 20-kilobase fragment (Fig. 2), indicating that neither the $^G\gamma$ nor the $^A\gamma$ gene is deleted in these DNAs. This result is especially significant in the case of Greek HPFH. Because the

chromosome bearing the Greek type of HPFH is thought to direct the synthesis of predominantly γ -globin chains⁷, it is possible that the γ gene is deleted in this syndrome. If this were the case, *BclI* digestion of DNA from a heterozygote would be expected to generate two fragments: one 20-kilobase fragment from the normal chromosome and another fragment of a different size from the HPFH chromosome. The abnormal fragment could be either longer or shorter than the normal 20 kilobases, depending on where the next *BclI* site resides in relation to the endpoint of the deletion. The presence of only one normal 20-kilobase *BclI* fragment in DNA from the Greek HPFH heterozygote strongly suggests that the γ gene is not deleted in this case. The possibility that the γ gene is deleted but that the new *BclI* site lies exactly 20 kilobases from the preserved *BclI* site (to the 3' side of the non-deleted γ gene) is unlikely from the results of other restriction endonuclease digestions. In particular, digestion of normal DNA with *BglII* yields a single 13.5-kilobase fragment containing both γ genes^{1,2,9}. DNA from our three cases of HPFH also yielded only a single 13.5-kilobase fragment when hybridized to γ probe (data not shown); in the case of Greek HPFH no additional band of a different size was seen that hybridized with the γ probe. The absence of a deletion of the γ gene in the Greek type of HPFH is further supported by the recent finding of small amounts of Hb F of the γ type in red cells of affected individuals¹⁰.

When *BglII*-digested DNA of our cases was hybridized with $\beta + \delta$ genomic DNA probes [from *PstI* subclones of H β G1 (ref. 11)], no bands were visible in the two cases of black HPFH (data not shown), confirming previous reports of total deletion of δ - and β -globin genes in these cases¹⁻³; in the case of Greek HPFH, a normal pattern was obtained, indicating the presence of normal-sized fragments containing the δ - and β -globin genes. As the Greek case is only heterozygous for HPFH, these normal bands are expected from the normal chromosome, so a deletion of δ and β loci from the Greek HPFH chromosome cannot be

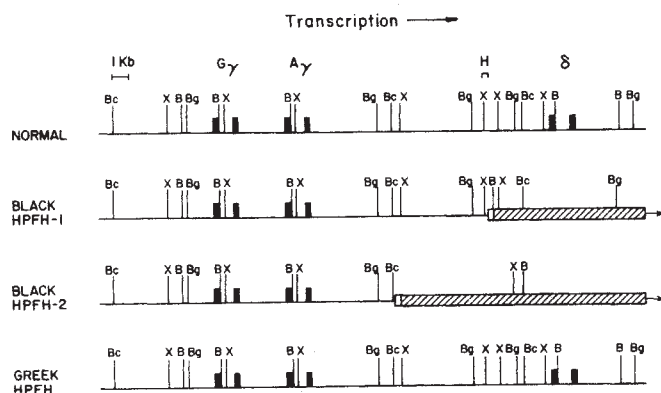


Fig. 1 Map of restriction endonuclease sites in and around the human non- α -globin genes in normal DNA and DNA from three different individuals with HPFH. The solid bars represent the coding sequences of the different non- α -globin genes, each interrupted by a large intervening sequence of non-coding DNA. The cross-hatched bars represent DNA to the 3' side of the deleted segments of DNA in HPFH DNA; the open portion of the bars indicates the region of uncertainty in the precise 5' end points of the deletions. Normal DNA was from the placenta of a black infant; black HPFH-1, DNA of a lymphoblastoid cell line derived from an American black with homozygous HPFH¹³; black HPFH-2, DNA of a lymphoblastoid cell line (GM2064) derived from a Ghanaian with homozygous HPFH^{3,14}; Greek HPFH, DNA of a lymphoblastoid cell line (GM2232) from a Greek HPFH heterozygote (case III-1 of family B in ref. 15). The experimental evidence for the assignment of the various restriction endonuclease sites around the γ -globin genes in normal DNA has been detailed elsewhere^{1,2,8,9}. The bracket indicates the position in normal DNA of H fragment DNA of clone H β G1 (ref. 11) that was used as a probe in some of the mapping experiments. Abbreviations: Bg, *BglII*; Bc, *BclI*; X, *XbaI*; B, *BamHI*; kb, kilobases.

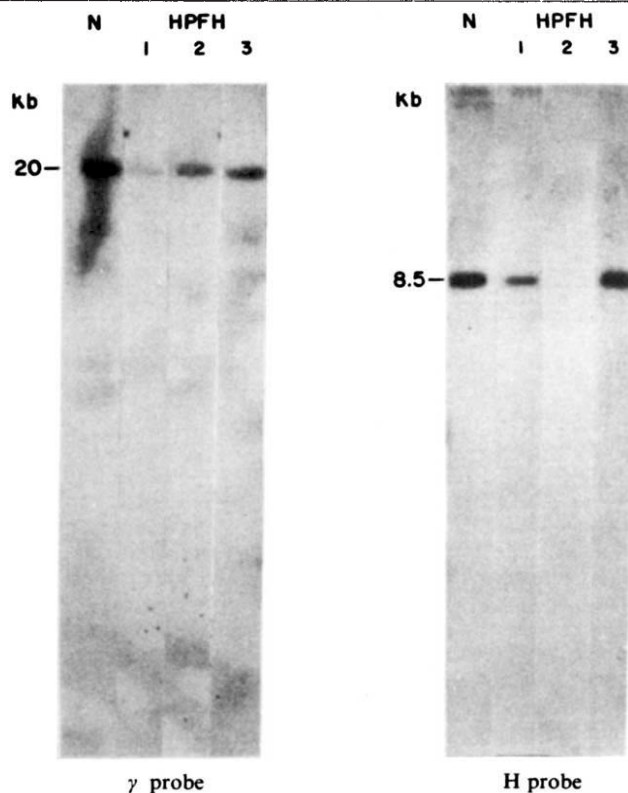


Fig. 2 Digestion of normal and HPFH DNAs with restriction endonuclease *BclI*. Digested DNA was fractionated by electrophoresis in a horizontal slab gel of 0.8% agarose as previously described², and the DNA was transferred to a nitrocellulose filter by the gel blotting technique of Southern¹⁶. The nitrocellulose filter was then hybridized in the presence of 10% dextran sulphate¹⁷ to the indicated cloned DNA probes labelled with ³²P by nick-translation as described elsewhere². The γ probe was a mixture of JW151 (a γ cDNA clone¹⁸) and pRP10 (an γ cDNA clone¹⁹). The H probe was an *EcoRI* subclone of H β G1 (ref. 11). Abbreviations: N, normal DNA; HPFH DNAs: 1, black HPFH-1; 2, black HPFH-2; 3, Greek HPFH.

ruled out on the basis of this result alone. However, if the δ and β loci were deleted on the affected chromosome in the Greek heterozygote, one would have expected to find an anomalous band in this DNA with one of the four enzymes and four probes that were tested (see Figs 3, 4); no anomalous bands were found and therefore a non- α -globin gene deletion could not be positively identified.

When *BclI*-digested DNA from our cases was hybridized to a probe from H fragment DNA [which is derived from a point 4.5 kilobases to the 5' side of the δ gene (Fig. 1 and refs 1, 11)], an 8.5 kilobase fragment was seen in all cases except for the Ghanaian case of black HPFH-2, in which no band was visible (Fig. 2). Thus, DNA corresponding to the region of the H fragment, in addition to δ - and β -globin gene DNA, is deleted in this case of black HPFH but not in the other.

To map the 5' end of the deletion in black HPFH-2, the enzyme *XbaI* was used. The *XbaI*-digested DNA of black HPFH-1 and Greek HPFH, when hybridized to γ -gene probe, exhibited a band pattern similar to that of normal DNA (Fig. 3): a 3.7-kilobase band derived from the 5' end of γ gene, a 5.2-kilobase fragment spanning the γ and α genes and a 7.5-kilobase fragment derived from the 3' end of the α gene^{1,2,9}. In *XbaI*-digested DNA of black HPFH-2, the 3.7- and 5.2-kilobase fragments are present as expected, but the 7.5-kilobase fragment is replaced by one of 15 kilobases (Fig. 3). Thus, the *XbaI* site present at the 3' end of the 7.5-kilobase fragment in normal DNA is missing in this case of HPFH. As this *XbaI* site lies approximately 1 kilobase to the 3' side of a *BclI* site

that is present in both normal and HPFH-2 DNA, the deletion in black HPFH-2 must start somewhere between this *BclI* site and the *XbaI* site to its 3' side. Comparing the 5' ends of the deletions in the two types of HPFH, an additional 5 kilobases of DNA is deleted in black HPFH-2 compared with black HPFH-1 (Fig. 1).

To examine whether the 3' endpoints of the deletions in the two cases of black HPFH also differed, the enzyme *Bam*HI was used. *Bam*HI cleaves normal DNA to generate an 18-kilobase fragment which spans the γ and δ genes and that hybridizes to both γ - and δ -gene probes^{1,2,8,9}. *Bam*HI-digested DNA of black HPFH-1 exhibits, instead, a 14-kilobase fragment (Fig. 4 and refs 1, 2) which hybridizes to γ -gene probe but not δ -gene probe (because the δ gene has been deleted). If the 3' endpoints of the deletions in both cases of black HPFH are the same, *Bam*HI digestion of HPFH-2 should produce a fragment 5 kilobases shorter than 14 kilobases. However, instead of a 9-kilobase fragment, a 16-kilobase fragment is generated in HPFH-2 (Fig. 4). This indicates that not only the 5' endpoints of the deletions in HPFH-1 and HPFH-2 are different but that their 3' endpoints also are different. The precise location of the 3' endpoints of the deletions cannot be mapped in relation to the DNA flanking the 3' extremity of the normal β -globin gene because of lack of appropriate probes for this region.

The above analyses demonstrate that at least two types of extensive DNA deletions differing in both their 5' and 3' endpoints are associated with the common $\gamma\gamma + \gamma\delta$ variety of HPFH that occurs in blacks. The deletion (HPFH-1) with the lesser amount of deleted inter- γ - δ DNA has been previously reported^{1,2}. The possible significance of this deletion with respect to the presence of putative γ -gene control elements situated between the γ - and δ -globin genes has also been discussed in relation to the different types of deletions detected in $\delta\beta$ -thalassaemia in which the γ -globin genes are not as maximally or uniformly expressed as in HPFH^{1-5,12}. The deletion of an additional 5 kilobases of inter- γ - δ -gene DNA in HPFH-2 over HPFH-1 had no significant effect on the HPFH phenotype. Despite the presence of two different types of HPFH deletion in blacks, it is remarkable that the three black phenotypic homozygotes who have thus far been studied are truly homozygous for one HPFH genotype or the other, and that neither is

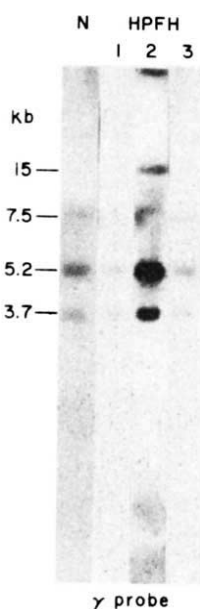


Fig. 3 Digestion of normal and HPFH DNAs with restriction endonuclease *XbaI*. Experimental procedure, γ -cDNA probe and abbreviations are the same as in Fig. 2.

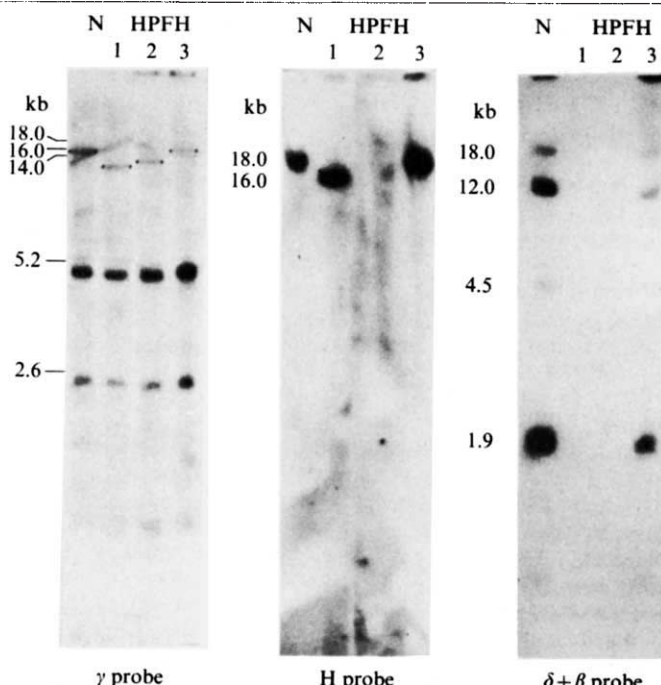


Fig. 4 Digestion of normal and HPFH DNAs with restriction endonuclease *Bam*HI. Experimental procedure, probes and abbreviations are the same as in Fig. 2. The $\delta + \beta$ probe was a mixture of H δ 1 and H β 1, *Pst*I subclones of H β G1 (ref. 11). The dots in the γ probe experiment indicate the positions of the faint 18-, 16- and 14-kilobase bands.

doubly heterozygous for both genotypes. This finding suggests some initial geographical or regional segregation of the two different mutations, with little subsequent mixing of the two different gene pools.

In the Greek $\gamma\gamma$ type of HPFH, no DNA deletion was detected; the putative γ -gene regulatory sequences in the inter- γ - δ DNA are therefore present. It is thus likely that the phenotype of Greek type HPFH is the result of a different molecular mechanism, such as a small DNA deletion or a base mutation in the $\gamma\gamma$ - $\gamma\delta$ gene region which has escaped detection by the techniques used in the present study.

We thank Drs A. Bank and A. E. Greene for providing the lymphoblastoid cell lines GM2064 and GM2232, Dr T. Maniatis and colleagues for the human δ - β gene clone H β G1 and its various subclones, Drs R. Poon and Y. W. Kan for the γ -cDNA clone pRP10, and Elaine Coupal for technical assistance. This work was supported in part by grants from the NSF, the National Foundation March of Dimes and the NIH.

Received 26 December 1979; accepted 14 March 1980.

1. Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Nature* **279**, 598-603 (1979).
2. Tuan, D., Biro, P. A., deRiel, J. K., Lazarus, H. & Forget, B. G. *Nucleic Acids Res.* **6**, 2519-2544 (1979).
3. Mears, J. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 1222-1226 (1978).
4. Ottolenghi, S. *et al. Nature* **278**, 654-657 (1979).
5. Bernards, R., Kooter, J. M. & Flavell, R. A. *Gene* **6**, 265-280 (1979).
6. Huisman, T. H. J. *et al. Ann. N.Y. Acad. Sci.* **232**, 107-124 (1974).
7. Huisman, T. H. J. *et al. J. clin. Invest.* **49**, 1035-1040 (1970).
8. Bernards, R., Little, P. F. R., Annison, G., Williamson, R. & Flavell, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4827-4831 (1979).
9. Little, P. F. R., Flavell, R. A., Kooter, J. M., Annison, G. & Williamson, R. *Nature* **278**, 227-231 (1979).
10. Clegg, J. B. *et al. Br. J. Haemat.* **43**, 521-536 (1979).
11. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G. & Maniatis, T. *Cell* **15**, 1157-1174 (1978).
12. Orkin, S. H., Alter, B. P. & Altay, C. J. *clin. Invest.* **64**, 866-869 (1979).
13. Forget, B. G. *et al. Cell* **7**, 323-329 (1976).
14. Ringelmann, B. *et al. Acta haemat.* **43**, 100-110 (1970).
15. Fessas, P. & Stamatoyannopoulos, G. *Blood* **24**, 223-240 (1964).
16. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
17. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
18. Wilson, J. T. *et al. Nucleic Acids Res.* **5**, 563-581 (1978).
19. Poon, R. P., Kan, Y. W. & Boyer, H. W. *Nucleic Acids Res.* **5**, 4625-4630 (1978).

Red cell sodium fluxes catalysed by the sodium pump in the absence of K^+ and ADP

Kenneth H. Lee

Department of Biochemistry, McGill University, Montreal, Quebec, Canada H3A 2T5

Rhoda Blostein*

Division of Hematology, Department of Medicine, Royal Victoria Hospital, 687 Pine Avenue West, Montreal, Quebec, Canada H3A 1A1

In the absence of extracellular Na^+ or K^+ , the sodium pump catalyses an ouabain-sensitive 'uncoupled' Na^+ efflux¹⁻⁴. With red cell ghosts Glynn and Karlisch⁵ showed that this Na^+ efflux is accompanied by ATP hydrolysis and that extracellular sodium ions, at low concentrations, inhibit this efflux as well as the associated ATP hydrolysis. At higher concentrations, extracellular sodium ions restore the hydrolysis of ATP^{3,6} but it is not known whether there is an associated increase in Na^+ efflux and, perhaps, an influx. To answer this question we have used inside-out red cell membrane vesicles which are specially suitable for controlling the composition of the medium at the two membrane surfaces while measuring $^{22}Na^+$ fluxes in both directions. We report here that the sodium pump can operate in a mode in which influx and efflux of sodium are associated with ATP hydrolysis. This mode is different from the Na - Na exchange described by Garrahan and Glynn⁷, and Glynn and Hoffman⁸, which requires ADP as well as ATP⁹ and is probably associated with ADP-ATP exchange rather than ATP hydrolysis^{10,11}.

For these experiments, we used inside-out membrane vesicles of human red cells and measured ATP-dependent $^{22}Na^+$ uptake and loss as described earlier¹² and detailed in Table 1. ATP-dependent $^{22}Na^+$ uptake was estimated from the increase in radioactivity retained in the vesicles after Millipore filtration; ATP-dependent loss of $^{22}Na^+$ was measured by the radioactivity appearing in the supernatant following removal of the vesicles by centrifugation, as it was not feasible to measure the decrease in vesicular $^{22}Na^+$ because the percentage loss was too small. Preliminary tests ascertained that the time intervals used (≤ 5 min for uptake and ≤ 15 min for loss) were sufficiently short that initial rates or near initial rates were being estimated.

We tested whether, in the absence of extracellular K^+ and without added ADP, an increase in extracellular Na^+ concentration is associated not only with inhibition of efflux at low extracellular Na^+ as described above, but also stimulation of efflux at higher extracellular Na^+ and whether this stimulation is associated with Na^+ influx. The results show (1) both inhibition and stimulation of efflux at low and high extracellular Na^+ , respectively; (2) Na^+ influx associated with the stimulated efflux; and (3) that these fluxes are paralleled by a decrease followed by an increase in the turnover of the phosphoenzyme intermediate of Na^+ -ATPase as the extracellular Na^+ is increased from zero to about 50 mM.

Because the vesicles are inside-out, the terms uptake and loss are opposite to those normally used, that is uptake and loss in vesicles are equivalent to the normal efflux and influx, respectively, in cells. Similarly, intravesicular Na^+ is equivalent to Na^+ normally at the extracellular surface and is designated Na_{ext} ; extravesicular Na^+ is Na^+ at the cytoplasmic surface and is designated Na_{cyl} .

The results given in Table 1 show that ATP-dependent 'uncoupled' Na^+ uptake (equivalent to normal efflux) is inhibited by a relatively low concentration of Na_{ext} (~ 5 mM) confirming the findings of others¹⁻⁴. This Na^+ flux is then increased at higher (~ 50 mM) Na_{ext} . The results show that the

increased uptake is associated with an ATP-dependent loss of $^{22}Na^+$ from $^{22}Na^+$ -loaded vesicles. A quantitative comparison of uptake and loss, even when measured concurrently, has not been attempted as it has not been technically feasible to measure both fluxes by the same method (see above). Although highly reproducible results were obtained for both, uptake values probably tended to be somewhat underestimated; this may be due to the removal of a portion of $^{22}Na^+$ from a competent but more leaky compartment, during washing of the vesicles on the filters. As shown in Table 1, the ATP-dependent $^{22}Na^+$ influx and efflux occur with only micromolar levels of ATP and in the absence of added ADP. In other experiments (not shown) we have ascertained that these fluxes, in particular $^{22}Na^+$ efflux (equivalent to normal influx), do not require ADP as well as ATP. Thus with 4 μ M ATP present, addition of ADP (5 μ M) did not affect the ATP-dependent $^{22}Na^+$ efflux. Furthermore, ATP-dependent Na^+ efflux is not decreased when either greater (fivefold greater) dilution of $^{22}Na^+$ -loaded vesicles is used to reduce the concentration of ADP formed due to ATP hydrolysis to extremely low levels (≤ 0.1 μ M at the end of the assay period, see Table 1 legend), or when creatine phosphate and creatine phosphokinase are included to regenerate ATP from ADP (not shown).

The results given in Table 2 show that ATP-dependent $^{22}Na^+$ loss (equivalent to normal influx) and uptake (equivalent to normal efflux) observed at relatively high Na_{ext} (~ 50 mM) have properties consistent with the conclusion that these Na^+ movements are manifestations of the activity of the ouabain-sensitive Na^+ pump. Thus: (1) both are ATP-stimulated increases in unidirectional fluxes; (2) both are sensitive to inhibition by ouabain as intravesicular ouabain (equivalent to extracellular ouabain) inhibits both activities; and (3) ATP-dependent $^{22}Na^+$ loss (equivalent to normal influx) is absolutely

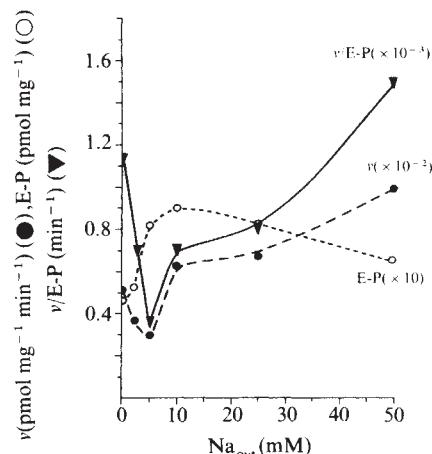


Fig. 1 Effect of Na_{ext} on Na^+ -ATPase and on the steady-state level and apparent turnover of phosphoenzyme. Vesicles (2.4 mg protein ml^{-1}) were equilibrated by diluting them 1:1 with 100 mM KCl (baseline activity) or chloride salts of Na^+ and choline to obtain the desired final concentration of Na_{ext} as described in Table 1. The reaction was then started by adding 180 μ l reaction medium containing [γ - ^{32}P]ATP and various amounts of NaCl to 20 μ l vesicles so that the final concentrations were: vesicle protein, 0.12 mg ml^{-1} , Na_{cyl} , 5 mM and [γ - ^{32}P]ATP, 0.1 μ M (27,000 c.p.m. $pmol^{-1}$). The reaction was terminated after 30 s with 2.3 ml 5% trichloroacetic acid containing 5 mM NaH_2PO_4 and 2.5 mM ATP. Aliquots (1.0 ml) were taken for the measurement of ^{32}P (Na^+ -ATPase, v) and 1.5 ml, for the measurement of E - ^{32}P as described previously¹². Values are means of triplicate determinations. The baseline values (71.4 $pmol$ mg^{-1} min^{-1} for v and 0.085 $pmol$ mg^{-1} for E - ^{32}P) obtained with KCl (Na^+ omitted) were subtracted. Six similar experiments were carried out. The maximal amount of ADP formed due to ATP hydrolysis may be calculated from the maximal total activity observed (baseline plus activity at 50 mM Na_{ext}) as follows: $(71.4 + 103.1) pmol$ mg^{-1} $min^{-1} \times 0.12$ mg $ml^{-1} \times 0.5$ min = 10.47 $pmol$ ml^{-1} .

* To whom correspondence should be addressed.

Table 1 Effects of intravesicular Na^+ (Na_{ext}) on ATP-dependent Na^+ transport in the absence of ADP and K^+

Intravesicular Na^+ conc. mM Na_{ext}	^{22}Na uptake (Expt 70) (pmol min $^{-1}$ mg $^{-1}$)			^{22}Na loss (Expt 108) (pmol min $^{-1}$ mg $^{-1}$)		
	+ATP	-ATP	ATP-dependent uptake	+ATP	-ATP	ATP-dependent loss
0	868 $\pm$ 36	574 $\pm$ 30	294 $\pm$ 28	none	none	none
5	688 $\pm$ 24	516 $\pm$ 4	172 $\pm$ 14	443 $\pm$ 8	405 $\pm$ 6	38 $\pm$ 6
50	1,100 $\pm$ 38	604 $\pm$ 20	496 $\pm$ 24	7,040 $\pm$ 86	6,053 $\pm$ 96	987 $\pm$ 74

Inside-out vesicles (~ 5.5 mg protein ml $^{-1}$) suspended in 40 mM Tris glycylglycine, pH 7.4, containing 0.1 mM MgCl_2 were loaded by diluting 1:1 with 2 mM MgCl_2 containing 100 mM chloride salts of choline chloride and NaCl to obtain the desired concentrations (0, 5 or 50 mM Na^+). For measurements of $^{22}\text{Na}^+$ loss, $^{22}\text{NaCl}$ was included (3,385 c.p.m. μl^{-1} suspension). Equilibration was then carried out for 18 h at 0 °C followed by 30 min at 37 °C. For $^{22}\text{Na}^+$ uptake measurements, 20- μl vesicles (5.4 mg ml $^{-1}$) were added to 180 μl $^{22}\text{Na}^+$ medium (specific activity 558 c.p.m. nmol $^{-1}$) containing 5 mM NaCl with or without ATP (final concentration, 3.6 μM ; Boehringer-Mannheim ATP, Na-form converted to Tris-form). After 5 min at 37 °C, $^{22}\text{Na}^+$ taken up was measured by Millipore filtration as described previously¹². For $^{22}\text{Na}^+$ loss measurements, the $^{22}\text{Na}^+$ -loaded vesicles were centrifuged at 10,000g for 10 min at 0 °C and the pellet was washed by resuspension and centrifugation three times in 10 ml ice-cold 5 mM NaCl in 45 mM choline chloride, 20 mM Tris-glycylglycine, pH 7.4, containing 1 mM MgCl_2 (solution A). The pellet was suspended in the same solution at the approximate original protein concentration and 10 μl (5.8 mg ml $^{-1}$) were then added to 90 μl solution A with and without 5.4 μM ATP. After 11 min at 37 °C, 90- μl aliquots were removed and added to 0.4 ml ice-cold solution A and within 2 min centrifuged for 10 min at 10,000g. Aliquots (0.4 ml) of the supernatant were taken for the determination of $^{22}\text{Na}^+$ radioactivity. Each value shown is the difference between two sets of quadruplicate determinations $\pm$ s.e.m., one without and one with ATP. The concentrations of extracellular Na^+ given are approximate, as after equilibration about 70–80% equilibration is actually achieved (other experiments, not shown). In similar experiments carried out at much higher dilution of vesicles (1/50 compared to 1/10) ATP-dependent $^{22}\text{Na}^+$ efflux was not decreased; at the higher dilution, the total ADP concentration was extremely low (0.06 μM at the end of the incubation as estimated by ^{32}P release from $[\gamma\text{-}^{32}\text{P}]\text{ATP}$).

Table 2 Some properties of ATP-dependent Na^+ transport in the absence of added ADP and K^+

Properties tested		Addition	$^{22}\text{Na}^+$ uptake (pmol min $^{-1}$ mg $^{-1}$) (ATP-dependent)	$^{22}\text{Na}^+$ loss (pmol min $^{-1}$ mg $^{-1}$) (ATP-dependent)
Expt 72-3	Ouabain sensitivity	−Ouabain (control)	984 ± 17	1,391 ± 300
		+Ouabain (50 μM intravesicular)	27 ± 20	255 ± 150
Expt 106	Dependence on Na _{cyt}	−Na _{cyt}	—	38 ± 150
		+Na _{cyt}	—	1,004 ± 145

Ouabain-containing vesicles were prepared by including 50 μM ouabain during the vesiculation process. Both the control and ouabain vesicles were loaded with 25 mM NaCl and assays were performed as described in Table 1. Na_{cyt} (5 mM) was either included or omitted as indicated. The ATP concentrations were 100 μM and 6 μM and the final protein concentrations, 0.45 mg ml $^{-1}$ and 0.65 mg ml $^{-1}$ in experiments 72-3 and 106, respectively.

dependent on the presence of Na^+ at the opposite side of the membrane, that is, cytoplasmic Na^+ . Other experiments (not shown) at varying concentrations of Na_{cyt} showed that ATP-dependent $^{22}\text{Na}^+$ loss was close to maximal with 5 mM Na_{cyt} . With ATP present (100 μM) we have detected stimulation of both $^{22}\text{Na}^+$ uptake and loss by addition of ADP, but only at ADP concentration ≥ 50 μM (three experiments, not shown). This is, presumably, the Na-Na exchange associated with ADP-ATP exchange.

To gain insight to the mechanism of the effects of extracellular Na^+ on Na^+ influx at low levels of ATP, we followed the apparent turnover of the phosphoenzyme intermediate of Na^+ -ATPase. For this purpose we measured the rate of Na^+ -ATPase, v , and the steady-state level of phosphoenzyme, $[\text{E-}^{32}\text{P}]$, using low concentrations (0.1 μM) of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ as described earlier¹³, to minimize nonspecific labelling. The apparent turnover of $\text{E-}^{32}\text{P}$ was calculated as the ratio $v/\text{E-}^{32}\text{P}$. In agreement with Glynn and Karlsh who used resealed red cell ghosts³, Na^+ -ATPase decreased as Na_{ext} increased from zero to about 5 mM and then increased as extracellular Na^+ was increased further to ~ 50 mM (Fig. 1).

Although the decrease in hydrolytic activity was often small or not always apparent (other experiments, not shown), dramatic changes in the apparent turnover of $\text{E-}^{32}\text{P}$ were always observed (Fig. 1). The decrease in turnover at low extracellular Na^+ , from 1,143 min $^{-1}$ to 367 min $^{-1}$, confirms a similar observation by Beaugé and Glynn¹¹ using a purified kidney preparation. They observed that Na^+ , at low concentrations, slowed the rate of phosphoenzyme breakdown and suggested that this effect is at inhibitory extracellular sites. Previous studies in our laboratory¹² have shown that cytoplasmic Na^+ catalyses the phosphorylation of the enzyme; the present finding that the apparent turnover of the phosphoenzyme is increased at relatively high concentrations of Na_{ext} is consistent with the idea that, in the

same conditions, extracellular Na^+ at high concentration stimulates the rate of dephosphorylation.

From this study, we conclude that in the presence of ATP at low concentration but absence of K^+ and added ADP, extracellular Na^+ at high concentrations is transported to the cytoplasmic compartment in exchange for cytoplasmic Na^+ , in a manner analogous to the exchange of K_{ext} for Na_{cyt} in normal conditions. Thus Na_{ext} at high concentrations stimulated dephosphorylation of the phosphoenzyme intermediate presumably in the reaction $\text{E}_2\text{-P} + \text{Na}_{\text{ext}} \rightarrow \text{Na} \cdot \text{E}_2 + \text{P}$; the resulting enzyme-Na complex may then undergo a conformational change, $\text{Na} \cdot \text{E}_2 \rightarrow \text{Na} \cdot \text{E}_1$ (compare $\text{K} \cdot \text{E}_2 \rightarrow \text{K} \cdot \text{E}_1$) before the release of Na^+ at the cytoplasmic side of the membrane. Note that in other experiments (not shown) we observed that ammonium ions, which are considered to be congeners of K^+ , stimulate ATP-dependent Na^+ uptake (equivalent to normal efflux) but inhibit the loss (equivalent to normal influx) consistent with the idea that they compete for extracellular Na^+ acting at the K^+ sites involved in dephosphorylation.

We thank Dr Rose M. Johnstone, Department of Biochemistry, McGill University, for helpful criticisms of the manuscript. The financial support of the MRC of Canada is gratefully acknowledged.

Received 3 January; accepted 5 March 1980.

- Garrahan, P. J. & Glynn, I. M. *J. Physiol., Lond.* **192**, 159–174 (1967).
- Garrahan, P. J. & Glynn, I. M. *J. Physiol., Lond.* **192**, 175–188 (1967).
- Glynn, I. M. *et al. Ann. N.Y. Acad. Sci.* **242**, 357–371 (1974).
- Lew, V. L., Hardy, M. A. & Ellory, J. C. *Biochim. biophys. Acta* **323**, 251–266 (1973).
- Glynn, I. M. & Karlsh, S. J. D. *J. Physiol., Lond.* **256**, 465–496 (1976).
- Beaugé, L. A. & Campillo, E. D. *Biochim. biophys. Acta* **433**, 547–554 (1976).
- Garrahan, P. J. & Glynn, I. M. *J. Physiol., Lond.* **192**, 189–210 (1967).
- Glynn, I. M. & Hoffman, J. F. *J. Physiol., Lond.* **218**, 239–256 (1971).
- Cavieser, J. D. & Glynn, I. M. *J. Physiol., Lond.* **263**, 214–215 (1976).
- Blostein, R. *J. biol. Chem.* **245**, 270–275 (1970).
- Beaugé, L. A. & Glynn, I. M. *J. Physiol., Lond.* **289**, 17–31 (1979).
- Blostein, R. *J. biol. Chem.* **254**, 6673–6677 (1979).
- Blostein, R. & Chu, L. *J. biol. Chem.* **252**, 3035–3043 (1977).

Evidence for mutation in an *I-A* gene

Ted H. Hansen*†, Roger W. Melvold†§, J. Scott Arn* & David H. Sachs*

* Transplantation Biology Section, Immunology Branch, DCBD, National Cancer Institute, NIH, Bethesda, Maryland 20205

† Department of Radiation Therapy, Harvard Medical School and Shields Warren Radiation Laboratory, New England Deaconess Hospital, 50 Binney Street, Boston, Massachusetts 02115

Present addresses: ‡Department of Immunology, Merck Sharp and Dohme Research Laboratories, Rahway, New Jersey 07065; §Medical Oncology Section, Northwestern University Cancer Center, Chicago Illinois 60611

Mutant mouse strains are important tools for immunogenetic studies of the regulation, structure and function of major histocompatibility (*H-2*) antigens¹⁻³. Several inbred strains have been established which carry *H-2*-linked mutations that cause changes in cell surface antigens as recognized by cytotoxic T cells and sometimes antibodies. Using intra-*H-2* recombinant haplotypes in skin graft complementation studies, lesions in several of these mutant strains have been genetically mapped to genes encoded in the *K*- or *D*-end of the *H-2* complex³. More precise mapping has often not been possible as many of the available mutant strains possessed lesions in the *K*-end of the *H-2^b* haplotype and a recombinant separating the *K^b* and *I-A^b* regions was not previously available. The genetic locations of the lesions in some of these mutant strains could therefore only be inferred when a serological and/or structural alteration was detected (see refs 1, 2). We describe here the use of a newly established recombinant strain, B10.MBR (ref. 4), for skin graft complementation studies to map *K*-end mutations. These studies provide the first genetic evidence that the B6.C-*H-2^{bm12}* mutation involves an *I-A^b* gene and also confirm that the lesions in two other mutants can be mapped to a gene in the *K^b* region.

The newly discovered inbred, congenic strain B10.MBR carries the recombinant haplotype *H-2^{bq1}* which juxtaposes *K^b* and *I-A^k* genetic markers (see ref. 4 and Table 1). Using B10.MBR mice, skin graft complementation studies were undertaken to map the mutations previously shown to involve the *K^b* and/or *I-A^b* genes. Three representative mutant strains were selected for this study, B6.C-*H-2^{bm1}*, B6-*H-2^{bm5}* and B6.C-*H-2^{bm12}*. The first two mutations are part of an extensive allelic series of mutations (more than 15) which have been mapped into the *K* to *I-A* interval⁵⁻⁷. The *H-2^{bm12}* mutation, while mapping also in *K* and/or *I-A*, is genetically complementary to the aforementioned allelic series, and was therefore presumed to involve a distinct second locus^{3,10}.

F₁ hybrids were produced by crossing mice of each of these mutant strains with B10.MBR mice. These hybrid mice then received skin grafts from B10 donors of the same sex as the recipients. As B10.MBR mice are congenic with B10 mice, only differences associated with *H-2* should be detected in this test system. If the lesion in the mutant haplotype being tested were restricted to a gene within the *K^b* region, then the hybrid should not reject the skin graft since B10.MBR provides *K^b* antigens in the recipient. However, if the lesion in the mutant

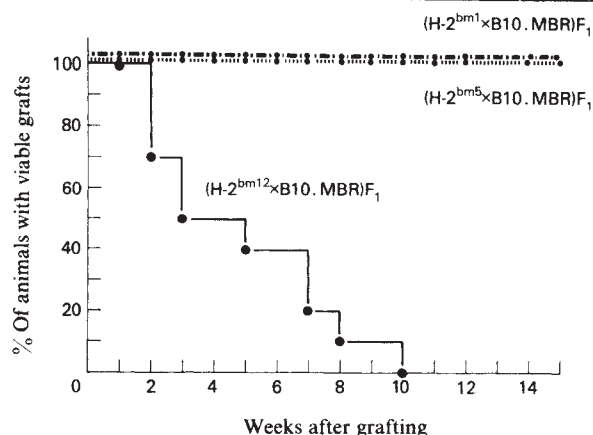


Fig. 1 Skin graft complementation study which maps the lesions in the *H-2^{bm1}*, *H-2^{bm5}* and *H-2^{bm12}* mutant haplotypes. Full thickness CSBL/10 tail skin grafts were trimmed to an approximate 0.5 × 1.0 cm size. Appropriately sized graft beds were prepared on the recipients' dorsal thorax by carefully removing skin with scissors without disturbing the panniculus carnosus. Five (*H-2^{bm1}* × B10.MBR)F₁, five (*H-2^{bm5}* × B10.MBR)F₁ and ten (*H-2^{bm12}* × B10.MBR)F₁ hybrid mice were skin grafted in this study. All recipients were males except for two in the last category (these females had rejection times of 2 and 3 weeks). The transplanted tissues were protected with a gauze dressing and a plaster bandage for 7 days. Grafts were scored as rejected when less than 10% of the donor tissue was viable by gross inspection.

being tested involved a gene to the right of the *K* 4 region, then the B10 skin graft would be rejected, as B10.MBR mice express *H-2^k* or *H-2^q* antigens in this interval.

The results of the skin graft complementation test of the *H-2^{bm1}* and *H-2^{bm5}* mutant strains are shown in Fig. 1. As indicated, the B10 skin grafts were not rejected by either the (*H-2^{bm1}* × B10.MBR)F₁ or the (*H-2^{bm5}* × B10.MBR)F₁ hybrid mice. These results provide definitive genetic evidence that the mutations in both of these strains occurred in a gene in the *K^b* region. In reference to the *H-2^{bm1}* haplotype (formerly referred to as *H-2^{bm1}* and *H-2^{bm2}*) the mapping to the *K^b* region substantiates serological^{7,9} and structural¹⁰ findings that the *H-2K^b* gene product in this mutant is different from the wild-type, *K^b* antigen. In addition to *H-2^{bm1}* and *H-2^{bm5}*, several other mutant strains are members of the same complementation group. Therefore, it is reasonable to infer that all members of this allelic series have mutations in a single gene located within the *K* region, presumably the *H-2K^b* locus which encodes a classical transplantation antigen.

In contrast to the findings with *H-2^{bm1}* and *H-2^{bm5}* mutations, the (*H-2^{bm12}* × B10.MBR)F₁ mice rejected B10 skin grafts (Fig. 1). This failure of the B10.MBR antigens to complement the *H-2^{bm12}* lesion indicated that this mutation involves a gene which maps to the right of the *K^b* region. Studies reported elsewhere determined that the lesion in the *H-2^{bm12}* haplotype mapped to the left of the *I-B* region (see Table 1). Taken together with the finding reported here, it can be concluded that the *H-2^{bm12}* mutation involves an *I-A^b* encoded molecule. Also this *I-A^b* encoded molecule can serve as a transplantation antigen.

Table 1 Summary of skin graft complementation data which map the lesions in the B6.C-*H-2^{bm1}* and B6.C-*H-2^{bm12}* strains

Donor	Recipient (mutant × recombinant)	<i>H-2</i> haplotype of recombinant strain								Fate of grafts	Conclusion regarding mapping of the mutations	Ref.
		<i>K</i>	<i>A</i>	<i>B</i>	<i>J</i>	<i>E</i>	<i>C</i>	<i>S</i>	<i>D</i>			
B10	B6.C- <i>H-2^{bm12}</i> × B10.A(4R)	k	k	b	b	b	b	b	b	Rejection	Maps to the left of <i>I-B^b</i>	8
B10	B6.C- <i>H-2^{bm1}</i> × B10.A(4R)	k	k	b	b	b	b	b	b	Rejection	Maps to the left of <i>I-B^b</i>	6
B6	B6.C- <i>H-2^{bm12}</i> × D2.GD	d	d	b	b	b	b	b	b	Rejection	Maps to the left of <i>I-B^b</i>	8
B6	B6.C- <i>H-2^{bm1}</i> × D2.GD	d	d	b	b	b	b	b	b	Rejection	Maps to the left of <i>I-B^b</i>	6
B10	B6.C- <i>H-2^{bm12}</i> × B10.MBR	b	k	k	k	k	k	k	q	Rejection	Maps to <i>I-A^b</i>	This paper
B10	B6.C- <i>H-2^{bm1}</i> × B10.MBR	b	k	k	k	k	k	k	q	Survival	Maps to <i>K^b</i>	This paper

As this is the first mutation shown to be in an *I*-region encoded histocompatibility antigen, it will facilitate several functional, structural and regulatory studies of this class of cell-surface molecules. Assuming that H-2^{bm12} is a simple mutation (that is, a single affected gene) studies using this mutant strain may elucidate the relationship between *I*-region encoded histocompatibility antigens and the serologically detected Ia antigens. Serological studies have determined that H-2^{bm12} cells express Ia antigens distinct from those found on wild-type, H-2^b cells (ref. 8 and T.H. *et al.*, in preparation). Functional studies of this *I*-region mutant are in progress to determine whether Ia antigens, *I*r gene products and *I*-region mixed lymphocyte reaction determinants are all different manifestations of the same cell-surface antigen.

Received 13 November 1979; accepted 12 March 1980.

1. Klein, J. *Adv. Immun.* **26**, 56–146 (1978).
2. McKenzie, I. F. C., Pang, T. & Blanden, R. V. *Immun. Rev.* **35**, 181–230 (1977).
3. Kohn, H. I. *et al. Immunogenetics* **7**, 279–294 (1978).
4. Sachs, D. H., Arn, J. S. & Hansen, T. H. *J. Immun.* **123**, 1965–1969 (1979).
5. Bailey, D. W., Snell, G. D. & Cherry, M. in *Proc. Symp. Immunogenetics of the H-2 System*, 155–162 (Karger, Basel, 1971).
6. McKenzie, I. F. C., Morgan, G. M., Melvold, R. W. & Kohn, H. I. *Immunogenetics* **3**, 241–251 (1976).
7. Apt, A. S. *et al. Immunogenetics* **1**, 444–451 (1975).
8. McKenzie, I. F. C. *et al. J. exp. Med.* (in the press).
9. Klein, J., Hauptfeld, M. & Hauptfeld, V. *J. exp. Med.* **140**, 1127–1131 (1974).
10. Brown, J. L. & Nathanson, S. G. *J. Immun.* **118**, 98–102 (1977).

Natural killer cells mediate lysis of embryonal carcinoma cells lacking MHC

Peter Stern, Magnus Gidlund, Anders Örn & Hans Wigzell

Department of Immunology, BMC, Box 582, Uppsala University, 75123 Uppsala, Sweden

Mouse teratocarcinomas are tumours of the germ cells¹ whose tumorigenicity is a property of a pluripotent stem cell type denoted the embryonal carcinoma (EC) cell. These cells seem to be analogous to cells of the early mouse embryo and will indeed contribute to the normal development of a mouse if injected into blastocysts². EC cells display an unusual transplantation behaviour and can frequently be grafted, giving tumours in several different mouse strains³. Thus, the protective immune mechanisms against such tumours may be less efficient than, and/or of a different nature from, those against most other tissues. Cytolytic T cells generally fail to kill EC cells^{4–6}, although there is one report of the apparent induction of specific anti-EC T-killer cells⁷. The failures^{4–6} are easily explained as being due to the absence of major histocompatibility complex (MHC) products on the surface of ECs^{3,8}. Recently, however, a new cell type, the natural killer (NK) cell, has been shown to have potent lytic activity when tested against several tumour targets^{9–11}. Targets of NK cells also include normal thymocytes¹² and certain stem cells of the bone marrow¹³. As NK cells seem to display a select but distinct specificity pattern from T lymphocytes and because NK targets include normal cells at an 'immature' stage of differentiation, we considered it possible that NK cells provide an alternative to T cells in producing cell-mediated immune reactions against teratocarcinomas. Here we report that murine NK cells are highly efficient cytolytic aggressors to EC cells whereas more differentiated cell lines derived from the latter are less susceptible or completely resistant.

Figure 1 illustrates that embryonal carcinoma cells are quite susceptible targets for normal spleen cells in short-term lytic assays; for example, the EC line Nulli-SCC.1 is almost as susceptible as the 'classical' NK target, YAC, a Moloney virus-induced lymphoma¹⁴. However, more differentiated endodermal lines derived from ECs were significantly more resistant to this lysis or even refractory, despite the fact that spleen cells from tilorone-treated mice were used as effector cells. This

Table 1 NK nature of effector cells lysing EC cells (Nulli-SCC.1 or PC 13) or YAC cells

Treatment	Target	Effector: target ratio			
		100:1	50:1	25:1	12.5:1
Control	YAC	21.0	11.1	5.7	4.5
Nylon wool		34.2	25.3	17.3	12.6
Complement (C)		26.7	16.7	10.2	5.3
Anti-Thy. 1 + C		27.9	19.3	13.2	5.9
Anti-asialo-GM ₁ + C'		3.2	1.4	0.8	0.2
Control	Nulli	12.2	7.2	6.5	7.0
Nylon wool		21.0	16.2	11.7	8.6
Complement (C)		16.8	10.2	7.5	5.1
Anti-Thy. 1 C'		15.7	9.5	6.4	6.7
Anti-asialo-GM ₁		1.0	2.3	0.6	0.5
Control	PC13	17.1	13.9	8.1	6.9
Nylon wool		18.2	15.7	11.4	8.8
Complement (C)		14.1	10.2	6.1	3.2
Anti-Thy. 1 + C'		16.8	10.1	6.7	6.7
Anti-asialo-GM ₁ + C'		3.2	2.7	2.9	2.7

Effector cells were normal spleen cells from 3-month-old CBA/H mice. The cells were treated as defined previously^{16,21}. Fresh guinea pig serum absorbed with agarose served as C source.

treatment, by a mechanism involving interferon, greatly enhances the lytic ability of NK cells¹⁵. Two additional EC lines (PSMB and OC15S1) were also studied and found to be highly sensitive to the lytic action of normal spleen cells (data not shown). Figure 2 confirms that the lysis of the EC cells is effected by NK cells. It is well established that NK levels are under genetic control, yielding 'high' and 'low' NK strains¹⁰. The ability

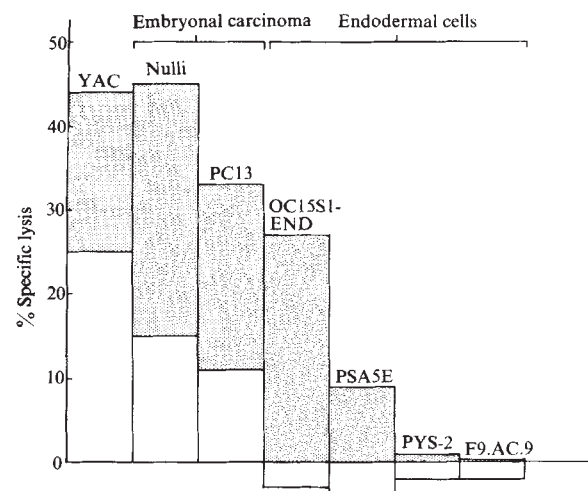


Fig. 1 NK lysis of YAC or various teratocarcinoma-derived cell lines using CBA/H spleen cells at an effector to target ratio of 100:1. Tilorone treatment and the 4-h ⁵¹Cr-release assay were done as described previously¹¹. The teratocarcinoma lines are as follows: PMSB is a pluripotent EC line and was grown as previously described where the stem cells do not differentiate and do not express H-2 antigens⁸. Nulli-SCC.1 is a line of homogeneous ECs which do not differentiate *in vivo* or *in vitro* and do not express H-2 (ref. 8). PC13 is a pluripotent EC line which remains undifferentiated when grown at high density *in vitro*; it will differentiate to endodermal cells in the presence of retinoic acid²⁰. OCC15S1, when subcultured at high density daily, maintains an EC phenotype²⁰; OCC15S1-END cells were obtained as the result of growth of the line at low density and there were only about 1% EC cells using anti-Forsman monoclonal antibodies in immunofluorescence tests²². The other EC lines were, by this criterion, as well as morphologically, homogeneous. The EC lines were routinely checked for H-2 and Ia determinants using specific antisera raised in congenic mice in a ¹²⁵I-protein A test²³. These cells do not express MHC products of the 129Sv(H-2^b) strain from which they are derived. The endodermal lines are all derivatives of 129 strain teratocarcinomas. PYS-2 is a parietal yolk sac carcinoma²⁴. PSA5E is an endodermal line derived from embryoid bodies grown *in vitro*²⁵. F9.AC.9 is cloned endoderm derived by retinoic acid treatment of F9 EC (ref. 26). Target cells were all grown in Dulbecco's modified Eagle's medium with 10% fetal calf serum (FCS) except YAC (Moloney lymphoma, A/Sn(H-2^a)) which was kept in RPMI-1640, 10% FCS. Teratocarcinoma cells were prepared as targets by brief trypsinization at 0.1% in EDTA-phosphate-buffered saline (PBS). Specific lysis in % = (c.p.m. in test - c.p.m. in medium control)/(c.p.m. in detergent - c.p.m. in medium control) × 100. Spontaneous release was never above 20% and there was less than 5% variation in triplicates. Open blocks denote lysis as caused using normal spleen cells whereas shaded areas indicate the increment in lysis obtained using spleen cells from tilorone-treated mice.

of normal spleen cells to lyse the EC cells has the typical genetic profile of NK cells against YAC cells. The relative activities of the same panel of effector cells against the other EC cell lines are consistent with these results. In contrast, the differentiated endodermal cells (for example, PYS-2, Fig. 2) are resistant to lysis by the same effector cells.

Any analysis of claimed NK activity must exclude the role of other potential effector cells. Table 1 shows fractionation experiments attempting to eliminate in a selective manner adherent cells, B and T lymphocytes or NK cells. Removal of macrophages and B cells using nylon wool columns resulted in an increased lytic activity of the remaining cells against the EC targets. Similarly, treatment of the effector cells using anti-Thy.1 antibodies and complement (C) failed to eliminate the killer cells for EC targets. Also, spleen cells from nude mice (devoid of T cells) were highly active against EC targets (data not shown), thus further excluding T lymphocytes as the effector cells involved. Finally, anti-asialoganglioside M₁ (GM₁) antisera in the presence of C (a procedure selective for NK cells¹⁶) abolished the lytic ability of the anti-EC effectors. This implicates in a positive manner murine NK cells as the major effector cells against the embryonal carcinoma cells. We thus conclude that mouse natural killer cells can efficiently kill murine EC cells (PSMB, OC15S1, Nulli-SCC.1 and PC13) whereas differentiated endodermal lines are either less sensitive (OC15S1-END, PSA5E) or completely resistant (PYS-2, F9.AC.9).

The precise specificity of NK cells is unknown and there is disagreement as to their heterogeneity¹⁰. As teratocarcinoma cells grow as well attached monolayers *in vitro*, they can be used as possible absorbents for the NK effector cells. Figure 3 depicts two separate experiments demonstrating the ability of EC cells (the NK-susceptible targets) to remove NK activity for YAC as well as for EC targets. In contrast, the NK-resistant endodermal cells failed to remove NK cells in any significant manner. The observed effects cannot be explained by detached tumour cells in the effector populations, as less than 0.1% of the unattached spleen cells could have been tumour cells as defined by light microscopy. Thus, a sizeable proportion of NK cells with lytic ability for YAC lymphoma cells must also express binding properties to EC but not endodermal cell lines. Use of this monolayer procedure may now allow an analysis of possible murine NK subgroup heterogeneity in the same manner as previously shown to be so productive in analysis of specific killer T lymphocytes¹⁷.

As stated above, cytotoxic T cells seem under most circumstances to be unable to kill embryonal carcinoma cells using normal chemically or virally modified cells⁴⁻⁶. There is one report of apparent T-cell killing of murine teratocarcinoma cells⁷, but we have been unable to obtain any significant immunity of T-cell type against the presently tested EC cells (to be published). In contrast, they can be seen by lytic NK cells in an efficient manner. Thus, we believe that EC cells, lacking MHC products,

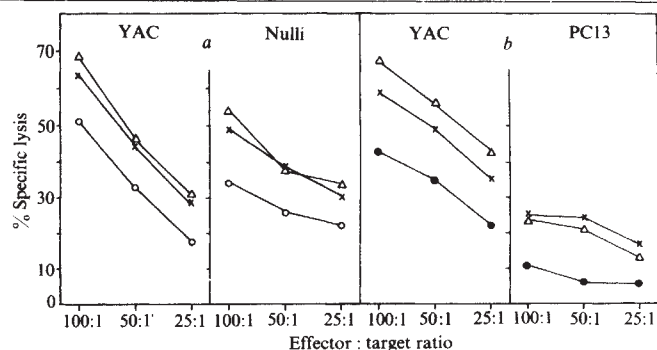


Fig. 3 Selective depletion of NK cells on NK-susceptible EC monolayers. Spleen cells from tilorone-treated mice were added at 3×10^7 cells per ml in 2 ml of RPMI-1640, 10 mM HEPES 5% FCS to 9-cm Petri dishes with no cells or with confluent monolayers of PYS-2 endodermal cells or PC13 or Nulli-SCC.1 EC cells. The cells were incubated at 37 °C in a CO₂ incubator for 45 min. The unattached cells were then recovered by pouring off the medium. Cell recovery was similar on all plates and in the range of 30%. Targets used were either YAC or Nulli-SCC.1 (a), or YAC or PC13 (b). a: ○, Nulli-SCC.1 monolayer-depleted effector cells; △, PYS-2 monolayer-depleted effectors; ×, normal Petri dish-depleted effectors. b, As for a except using PC13 monolayer-depleted effectors ●.

are suitable targets for distinguishing NK- and T-cell killing. In this respect, we consider it likely that recent attempts to map immune response genes or transplantation loci providing relative *in vivo* resistance to transplanted teratocarcinomas with suggested linkage to H-2 (refs 18, 19) were, in fact, measuring regulatory genes for NK levels which exhibit a similar linkage¹⁰.

In conclusion, teratocarcinoma-derived cells of the most primitive type, EC, have been found to be highly susceptible to murine NK cells whereas differentiated cell lines from the same tumours are markedly more resistant. Preliminary results show that if EC cells are induced to differentiate under the influence of retinoic acid²⁰ there is a parallel loss of NK sensitivity. Molecular analysis of this process may result in a more refined concept of why NK cells have a specificity pattern for certain malignant cells as well as groups of relatively immature normal cell types. One attractive possibility is that NK cells recognize embryonic structures which are lost on further differentiation, but reappear during malignant transformation.

We thank Dr M. Kasai for the anti-asialo-GM₁ antibodies, and C. Graham for PC13, OCCISI and F9.AC.9. This work was supported by an EMBO fellowship (P.S.), a contract from the NIH, and the Swedish Cancer Society.

Received 31 December 1979; accepted 12 March 1980.

- Martin, G. R. *Cell* **5**, 229-243 (1975).
- Mintz, B., Illmense, K. & Gearhart, J. D. in *Roche Symp. on Teratomas and Differentiation* (eds Sherman, M. & Solter, D.) 45 (Academic, New York, 1975).
- Arlitz, K. & Jacob, F. *Transplantation* **17**, 632-634 (1974).
- Forman, J. & Vitetta, E. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3661-3665 (1975).
- Zinkernagel, R. M. & Oldstone, M. B. O. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3666-3670 (1976).
- Doherty, P. C., Solter, D. & Knowles, B. B. *Nature* **266**, 361-362 (1977).
- Wagner, H., Starzinski-Powitz, A., Rollinghoff, M., Golstein, P. & Jakob, M. J. *J. exp. Med.* **147**, 251-264 (1978).
- Stern, P., Martin, G. R. & Evans, M. J. *Cell* **6**, 455-465 (1975).
- Herberman, R. B., Nunn, M. E., Holden, H. T. & Lavrin, D. H. *Int. J. Cancer* **16**, 230-239 (1975).
- Kiessling, R. & Wigzell, H. *Immun. Rev.* **44**, 165-208 (1979).
- Kiessling, R., Klein, E. & Wigzell, H. *Eur. J. Immun.* **5**, 112-117 (1975).
- Hansson, M., Kiessling, R., Andersson, B., Kärre, K. & Roder, J. *Nature* **278**, 174-176 (1979).
- Kiessling, R. *et al. Eur. J. Immun.* **7**, 655-660 (1977).
- Herberman, R. B. & Holden, H. T. *Adv. Cancer Res.* **27**, 305-377 (1978).
- Gidlund, M., Örn, A., Wigzell, H., Senik, A. & Gresser, I. *Nature* **273**, 759-761 (1978).
- Kasai, M., Iwamori, M., Nagai, Y., Okumura, K. & Tada, T. *J. Immun.* (in the press).
- Brondz, B. D. *Transplant. Rev.* **10**, 112-151 (1972).
- Avner, P. R. *et al. Immunogenetics* **7**, 103-115 (1978).
- Siegler, E. L., Tick, N., Teresky, A. K., Rosenstrauss, M. & Levine, A. J. *Immunogenetics* **9**, 207-220 (1979).
- Adamson, E. D., Gaunt, S. J. & Graham, C. F. *Cell* **17**, 469-476 (1979).
- Kiessling, R. *et al. J. exp. Med.* **143**, 772-780 (1976).
- Stern, P. L. *Cell* **14**, 775-783 (1978).
- Welsh, K. I., Dorval, G. & Wigzell, H. *Nature* **254**, 67 (1975).
- Lehman, J. M., Speers, W. C., Swartzendruber, D. E. & Pierce, G. B. *J. Cell. Physiol.* **84**, 13-28 (1974).
- Adamson, E. D., Evans, M. J. & Magrane, G. G. *Eur. J. Biochem.* **79**, 607-615 (1977).
- Solter, D., Shevinsky, L., Knowles, B. B. & Strickland, S. *Dev. Biol.* **70**, 515-521 (1979).

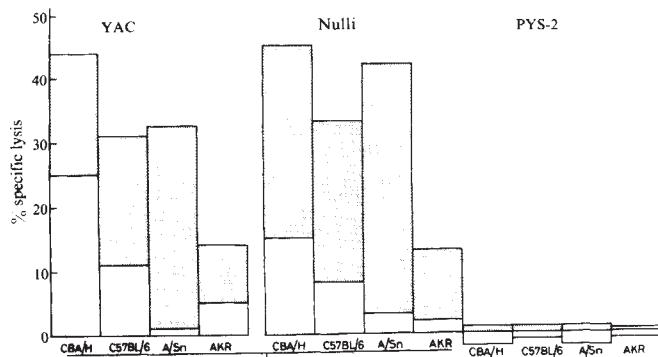


Fig. 2 Comparison of genetic distribution of NK levels as measured against the 'classical' NK target YAC and the distribution profile of lytic effector cells against EC cells (Nulli-SCC.1) and the insensitive endodermal line PYS-2. Effector:target ratio 100:1. With or without tilorone treatment of spleen cell donors as in Fig. 1.

BOOK REVIEWS

Evolution, theology and the Victorian scientist

Stephen Jay Gould

UNFORTUNATELY, the important issue of evolution and its historical relation with theology has been badly distorted by the stereotype of crusading, modern hero versus pompous, superannuated windbag — Huxley and Wilberforce for the British, Darrow and Bryan for Americans. Yet anyone who has ever read the leading theological scientists — from Paley whom Darwin read with pleasure in his youth, to Sedgwick who taught him geology and later branded his great work as crass materialism, to Mivart whose critique inspired the only new chapter added to later editions of the *Origin* — knows full well that the actual debates could scarcely have been further from the image of bigoted Bible-thumping versus enlightened rationalism. For these men were not miracle mongers, asserting God's continuous intervention against the lawful regularities of science; they were not, despite Lyell's words, trying "to cut rather than patiently to untie the Gordian knot". To be sure, they found God manifest in His works (nature), but their tool was science, their commitment to regularity of law and 'knowability'.

Equally unfortunately, this striking discrepancy between stereotype and actuality has engendered the equally misleading myth that there never really was any conflict between science and religion, that the theme embodied in A. D. White's famous treatise, *A History of the Warfare of Science with Theology in Christendom*, is at worst a put-on, at best a *Scheinproblem*. It is the great merit of Gillespie's book that he steers the proper path between these fallacious readings and manages to explain so well how working scientists framed the issue of conflict between science and theology in Darwin's day.

Borrowing a phrase and concept from the most fashionable intellectual of the moment, the French philosopher Michel Foucault, Gillespie recasts the debate as a

Charles Darwin and the Problem of Creation. By Neal C. Gillespie. Pp.224. (University of Chicago Press: Chicago and London, 1979.) £11.55.

contrast between two "epistemes" or world views — creationism and positivism. (Foucault, like T. S. Kuhn, tends to view intellectual history as a set of rapid switches between contrasting world views, each sufficiently encompassing to form a system necessarily excluding its rivals.) Creationism and positivism do not represent religion versus science, but two styles of doing science. (Many positivists, including Darwin at one stage of his life, believed in a personal God who may once have made the world, but declined to muck about with it thereafter.) Positivists insisted upon a fully natural and knowable system of causes, based upon unaltered laws manifest in processes now acting on Earth. The positivist, Gillespie claims, "limited scientific knowledge, which he saw as the only valid form of knowledge, to the laws of nature and to processes involving 'secondary', or natural causes exclusively". Only this kind of knowledge could be called science; indeed, only this kind of knowledge could be called valid at all. Creationists believed that an understanding of nature would reveal the working of God's mind. "The creationist", Gillespie argues, "saw the world and everything in it as being the result of direct or indirect divine activity. His science was inseparable from his theology".

Ironically, the positivist creed triumphed largely because it limited the scope of legitimate questions that science could ask about nature. No ultimate meanings, cosmic purposes or universal designs. Yet, in eschewing such grand themes, positivism gained precision and knowability — its questions could be answered and its answers could win general agreement. The

issue was not science versus God, but the range of legitimate questions within science. (The nineteenth century, after all, was not an age of rampant atheism; few positivists doubted the existence of a personal God.) Positivism held that science could do less, but could answer with agreement and authority. The victory of positivism marked the triumph of fruitfulness over interminable wrangling about the unanswerable.

Thus, the pre-Darwinian creationists — the true 'scientific creationists', in contrast with the fundamentalist imposters now parading under that banner in the United States — did believe that each new species arose *ab nihilo*, but did not use this credo to argue for God against science, miracles against law, or caprice against uniformity. Instead, they searched for regularity in the pattern of creation, hoping to specify its laws, if not its mechanism. And they held an (admittedly vague) belief in the existence of some law-like mode by which the Creator might manufacture his products. Yet, despite the brave words, their actual programme of research held out no hope for specifying or understanding such laws; they bogged down in the unanswerable. The recognition of this dead-end helps to explain why so many creationists were willing to accept Darwin's argument (but to recast it in more congenial terms as the belief that evolution represented God's mode of ordaining the history of life). As Charles Lyell wrote: "It would be more natural to suppose an ass to give rise to a striped offspring with the other characters of a zebra than that a zebra should come into being out of nothing".

But if the creationist episteme entered Darwin's age via evolution as God's mode, then positivists like Darwin and Huxley could rightly ask what creationism had added to their inquiry beyond obfuscation. How did the statement that God worked via evolution help anyone to discover the

mechanisms and processes? In this sense, God had become Laplace's unnecessary hypothesis. Worse than that. Darwin's own mechanism of natural selection seemed so purposeless and heartless to creationists that they sought God's characteristic (anthropomorphic) mode of action in such putative phenomena as variation intrinsically directed towards improvement or saltational origin of fully adapted higher forms of life. These beliefs emasculated natural selection by stripping away its essence — its creative role in building new forms by accumulating small, random, favourable variants — and relegated it to the negative role of executioner for the unfit. Again, the remnants of the creationist episteme became the source of a debate *within* science: what role did natural selection play in the origin of species?

Theology did not only invest science in the form of a lingering creationist episteme. It also influenced the outlook of positivists as well. Darwin became an agnostic in his later life and never (as a positivist) allowed his personal God any direct role in the daily operation of the Universe. But the *Origin* is studded with Darwin's theology; how could any great nineteenth-century work fail to record its influence? Gillespie argues cogently, for example, that the principle of natural selection itself can be viewed, in part, as Darwin's solution to the theological problem of evil. Buckland, and other creationists, had argued that predation was God's gift to prey because it ended their life swiftly before the pain of decrepitude set in. Darwin could not believe that a humane God would work in such a manner, especially when he considered such lingering forms of death as the living, but paralysed, caterpillar devoured by the larvae of wasps. Nature had its cruelties because God played no direct role. Darwin's concern with the problem of evil is evident in what may be the most striking, literary line of the *Origin*:

We behold the face of nature bright with gladness . . . We do not see, or we forget, that the birds which are idly singing round us mostly live on insects or seeds, and are thus constantly destroying life; or we forget how largely these songsters, or their eggs, or their nestlings, are destroyed by birds and beasts of prey.

Gillespie writes about these matters with commendable clarity and authority but without much verve. In many places, the book is little more than a series of quotations strung together with a minimum of connecting text — as if it had been compiled in the old style from index cards, one quote to the card and properly shuffled. One shouldn't condemn a colleague for excess of scholarly zeal, and a little learning may be a dangerous thing. But judicious selection larded with incisive commentary has its place as well, and *Parsifal* might have been even better an hour or two shorter.

More importantly, though I think that the conflict of epistemes is a better model for scientific change than old myths of inductive progress, I am not sure that it always serves Gillespie well. Although he admits that aspects of both epistemes can reside in single individuals, Gillespie does have a tendency to place scientists into one or the other camp. Creationists, pursuing old and unfruitful ways, therefore achieve the unconscious status of impeders. We assume, for example, that Mivart, whom Gillespie describes as dishonest or even odious (to Darwin at least), upheld old ways as a result of his theological bent. This may be an accurate description of his psychology — of his deeper motivation for attacking Darwinism in favour of quasi-theological, directed evolution. But his specific argument still has a logic that needs to be analysed in its own terms. Often, the logic is both cogent and potent. Mivart, for example, penned a brilliant (but forgotten) chapter on homology in *The Genesis of Species*. He argued that the development of serially homologous structures in single individuals obviously cannot be attributed to natural selection, but to "laws of growth". (The laws may be a product of selection, but this is a different matter.) He then claimed that homologous structures in related species often displayed the same morphological sequences as the serially homologous structures of a single individual. How then can homologies between species be attributed only to

natural selection acting upon plastic organic matter? Selection must be strongly constrained in its potential product by the same laws of growth that produce serial homologues in individuals.

Constraint of possible directions for evolutionary change by inherited patterns of development is just now emerging as an important theme in evolutionary biology; it also runs counter to the spirit of NeoDarwinian orthodoxy which exalted the power of selection and downgraded the constraints and consequences of development. Yet, if Mivart must bear the label of "creationist episteme", we may downgrade his message because we recognize its source.

As a scientist who also doubles as an amateur historian, I do appreciate the argument of historical colleagues that present status is an utterly inappropriate category for understanding past ideas, but I also believe that many rejected thinkers of the past appreciated great truths long forgotten by modern traditions. We must separate the logic of their argument from its psychology and genesis. I doubt that Gillespie will disagree, for scientists and historians approach the past both for different reasons and (or so it should be) with a common goal of appreciation in the broadest sense. □

Stephen Jay Gould is Professor of Geology at the Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts.

Perception model

Patrick Rabbitt

Decision Processes in Visual Perception. By D. Vickers. Pp. 406. (Academic: London and New York, 1979.) £23.80, \$55.

PSYCHOPHYSICISTS investigate regularities in human sensory discrimination in order to propose mathematical models for sense organs and neural structures which cannot, yet, be more directly explored in other ways. Their discipline demands a high level of ability, some mathematical sophistication, a detailed knowledge of a very large body of literature in physiology and psychology, and originality in the design and interpretation of experiments. Their work is not easily followed by most academic psychologists and is not followed at all by the public at large.

Douglas Vickers gives a novel and thoughtful account of the subject and attempts, as far as possible, to make it accessible to the general reader. The general reader must be patient and hard-working. Vickers believes that models for human brain function should be guided by what we can guess about the probable evolution of sense receptors and of their analytic backup in the central nervous

system. "Just as the ontogeny of an individual tends to recapitulate the evolution of the species . . . so the human visual system may be expected to embody a succession of progressively refined response sub-systems from [systems for] the most primitive discriminations up to the most complex and specific recognition . . ." (p. 17). This idea of the parallel, separate operation of series of separate systems of different degrees of complexity and subsuming different kinds of discrimination is a profitable one. It is not easy, otherwise, to make sense of recent results obtained from patients with occipital brain lesions who can make discriminations without, apparently, being 'aware' of the data on which they base their judgements. From the minimum perceptual requirements of a hypothetical primitive organism, Vickers attempts to deduce the various problems of discrimination and analysis of information which must arise, and to suggest the most probable mathematical and statistical techniques which the brain uses to solve them. He repeatedly returns to his main thesis that the separate component sub-systems jointly form "a kind of loose society of similar, though independent self-regulating units rather than a unified system of interdependent specialized functions such as are embodied in the

ordered sub-routines of most computer programs" (p. 17).

Vickers is a pleasant and learned guide to the history of his subject since the publication of Fechner's *Element der Psychophysik* in 1860 and his enthusiasm for historical detail and quotation is engaging. However this would be a poorer and less coherent book if he had not also attempted to offer a unifying view of his own in terms of his (new, updated, comprehensive) "accumulator model" for decision processes. Earlier versions of this model have been useful guides to data analysis and interesting sources of ideas for experiments. It is instructive to follow Vickers in his attempts to stretch the new comprehensive model beyond the statistical sampling of evidence and confidence levels for decisions, beyond even discriminations between two or more stimuli, choice reaction times and signal detection, to cover adaptation levels. With less

success he interprets some issues in prolonged vigilance and inspection (pp. 230-233), as well as an extremely difficult range of problems in perceptual organization (pp. 291-341), finally covering "Choice, Control and Consciousness" (pp. 359-370).

Of course the model clearly fails to cope with the massive tasks which Vickers sets it. Success is not possible in the current state of the art and Vickers is to be congratulated for subjecting his model to the widest possible range of tests — including those that it is bound to fail. There is no better way to proceed since it is precisely the points at which a model fails that are of compelling interest for further work. We can know more clearly what the problems are only when we realize what we still need to know before we can begin to solve them.

Patrick Rabbitt is a Lecturer in Experimental Psychology at the University of Oxford, UK.

Nonlinear optics in gases

Peter Knight

Nonlinear Optics of Free Atoms and Molecules. By D.C. Hanna, M.A. Yuratch and D. Cotter. Pp.351. (Springer: Heidelberg and New York, 1979.) DM 79, \$44.30.

THE nonlinear optical response of matter to intense radiation has been a major research field since the advent of the laser. An early spectacular demonstration of such a response was the frequency doubling of ruby laser light in a quartz crystal. Nonlinear optical processes in solids suffer from low efficiencies and damage problems. The pioneering observations of nonlinear optical processes in atomic and molecular gases by New and Ward freed the experimentalist from damage problems, and the field was further advanced by the proposal by Miles and Harris that a large enhancement in efficiency is possible when atomic or molecular resonances are exploited. Tunable frequency converters based upon nonlinear optical properties of gases are now available throughout the spectrum from the X-UV to the far infrared, using harmonic generation, stimulated electronic Raman and hyper-Raman scattering and related techniques. The authors of this book are well known for their extensive theoretical and experimental studies of nonlinear optics.

The emphasis in this much-needed book is upon the use of near-resonant nonlinear optical techniques to produce tunable radiation by frequency conversion. The first two chapters contain a careful discussion of the theory of nonlinear susceptibilities and radiation propagation. This is followed by chapters on harmonic

generation, stimulated electronic Raman scattering and hyper-Raman scattering, four-wave mixing, the nonlinear optical properties of molecules, and a final chapter on new and potentially useful developments (e.g. radiative collisions and phase conjugation). Theoretical developments are carefully integrated with experimental results, with valuable statements of what remains to be understood. Practical considerations (such as heat-pipe design) are discussed at length, as are theoretical complications of selection rules and angular momentum algebra. I particularly liked the careful attention paid by the authors to units and definitions.

There are a few things I didn't care for in this book. The idea of dressed, or 'adiabatic' atom-plus-field states is presented but is applied quite incorrectly in their discussion of optical resonant saturation. If dressed states are used to describe strong-field resonant excitation, it is quite wrong to state that such a saturating field stimulates transitions between dressed states as the authors do in Fig. 2.4 and elsewhere. I found their discussion of such concepts confused and misleading. It would also have added to the value of the book if an elementary derivation of the susceptibility formulae had been given from perturbation theory rather than relying on a reference to Butcher's standard but virtually unobtainable monograph.

This book is valuable despite its minor faults. I am sure it will become a standard handbook for physicists and chemists interested in the generation of tunable radiation by exploiting the nonlinear optical properties of atomic and molecular gases. □

Peter Knight is an SRC Advanced Fellow in the Optics Section, Imperial College, University of London, UK.

Russian pioneer in electricity and magnetism

I.B. Cohen

F.U.T. Aepinus: Essay on the Theory of Electricity and Magnetism. Translation by P. J. Connor. Introductory monograph and notes by R. W. Home. Pp.514. (Princeton University Press: Princeton, New Jersey, 1979.) £20.70.

FRANZ Ulrich Theodor Aepinus was an outstanding physicist of the eighteenth century. His *Tentamen Theoriae Electricitatis et Magnetismi*, published in St Petersburg in 1759, was a pioneering effort to introduce mathematical analysis into these two branches of physics. Despite its importance as an intellectual landmark, however, Aepinus's *Tentamen* was never widely read. The edition was small (650 copies); published in Russia, copies were not too easily obtained in European capitals. Furthermore, the book was not reprinted or translated until a Russian edition appeared in 1951, under the editorship of Ya. G. Dorfman. Most scientists who heard of Aepinus's work encountered it by way of a secondary source, the most important of which was the epitome of Aepinus's essay published in French in Paris in 1787 by the Abbé René-Just Haüy, the great pioneer crystallographer. The mathematician Gaspard Monge, founder of projective geometry, had earlier made a summary of the *Tentamen* for his own use. In 1801, Haüy's epitome appeared in a German translation. The first major account of Aepinus's work to be published in English was in the two-volume supplement to the third edition of the *Encyclopaedia Britannica*, where John Robinson included lengthy summaries of Aepinus's theories in his articles on electricity and on magnetism. These were given wider circulation when they were later reprinted in their entirety in Vol. 4 of Robinson's *System of Mechanical Philosophy*. Thomas Young included these theories in his *Course of Lectures on Natural Philosophy and the Mechanical Arts*.

Aepinus's theory of electricity was a development from Benjamin Franklin's theory in which "the American's incompletely worked out ideas were for the first time rendered precise and fully consistent with each other". Aepinus's first contribution to electricity was his discovery in 1756 that the reason a warmed tourmalin crystal attracts dust or ashes is electrical: he soon showed that this was different from the normal electrification obtained by friction, since no rubbing is required. Additionally, the warmed crystal did not simply acquire an overall charge, but rather gained "opposite electrical charges on two opposite faces". This

discovery was at once confirmed and applauded by others, including Benjamin Franklin.

Aepinus's work on tourmalin had been done in Rostock; soon thereafter he moved to St Petersburg and spent the remainder of his scientific life in the service of the Imperial Russian Academy of Sciences. In his book he remedied the following defects in the Franklinian theory of electricity: he got rid of Franklin's doctrine of electric "atmospheres", replacing it by the concept of electrical force acting at a distance; he explained how negatively charged bodies might attract one another by postulating that "common matter" deprived of its "normal" amount

of electrical fluid will repel similar "common matter"; finally, he attempted to reduce electrical science to a set of mathematical principles, an endeavour that fell somewhat short of the mark since Aepinus did not know the inverse-square law of electrical (and of magnetic) attraction. He also constructed a parallel theory for magnetic forces and effects.

This first translation of Aepinus's essay, by P. J. Connor of the Department of Classics of the University of Melbourne, reads well and seems to be accurate. The introductory monograph, by R. W. Home of the same university, is a mine of information concerning the life and career of Aepinus, his influence, the state of

electrical and magnetic experiment and theory in the eighteenth century, the organization of research in the Imperial Russian Academy of Sciences and much else. This monograph, it should be noted, is of book-length, just about as long as Aepinus's *Tentamen*, which it introduces.

The monograph and translation form a first-rate contribution to our knowledge of the history of science. Together they should help restore Aepinus to his proper place in the development of electricity and the use of mathematics in physical science.

I.B. Cohen is Professor in the Department of the History of Science, Harvard University, Cambridge, Massachusetts.

Birds' sense of direction

John R. Krebs

Avian Orientation and Navigation. By K. Schmidt-Koenig. Pp.180. (Academic: London and New York, 1979.) £13.80, \$32.

THIS is a well-written, balanced and accurate introduction to the literature on bird navigation and orientation up to the end of 1978. The major conclusions from work on pigeons are that birds find their way by using compasses based on the Sun and the Earth's magnetic field (nocturnal migrants also use the stars) in conjunction with an unidentified map. The map might

well be based on magnetic information acquired by the pigeon during its outward journey to the release site, since pigeons transported in an altered magnetic field show less accurate homeward orientation when released. The suggestion by Italian workers that pigeons may sniff their way home by using a scent map is still controversial. One of the most convincing experiments was one in which pigeons in their home loft were exposed to clockwise or counterclockwise deflected winds. When released, these birds flew with a predicted deviation from the correct homeward direction. However, the same effect is shown by pigeons with anaesthetized olfactory epithelia, suggesting that information other than scent must be related to wind direction.

One remarkable outcome of studies of homing is that pigeons have been shown to possess unsuspected sensory capabilities. Conditioning techniques have been used to show that pigeons can detect small changes in atmospheric pressure, infrasounds as low as 0.06 Hz, ultraviolet light and the plane of polarized light. The demonstration of ultraviolet sensitivity is an interesting case history. Previous workers had noted that the sensitivity of pigeons decreased towards wavelengths of 400 nm. However, below 400 nm sensitivity suddenly increases again to a second peak at 360–325 nm. We still do not know what role these various sensory abilities play in navigation and orientation, but infrasound might be a good candidate for a large-scale auditory map, since very low frequency sounds generated by waves, wind in mountain ranges and so on may carry for several thousand kilometres. The discovery that birds, in common with many other animals, can detect the Earth's magnetic field is also the result of orientation studies. Schmidt-Koenig's book was completed just before the recent demonstration by Walcott *et al.* (*Science* 205, 1027–1029; 1979) that pigeons have tiny crystals of magnetite in their heads which could be part of the sense organ used to detect magnetic fields.

As with homing pigeons, more is known about how natural migrants use compasses than about their maps. Even relatively complex migration routes may be followed entirely with a compass. Garden warblers (*Sylvia borin*), for example, seem to have a built-in clock and compass which tells them to fly SW for one month in the autumn, followed by SSE for another month. This takes the birds from central Europe to southern Spain and thence to west or central Africa. Whether such birds have any map sense remains uncertain.

I have summarized a small sample of the results described by Schmidt-Koenig and I recommend that anyone interested in the subject should read his book. □

John R. Krebs is a Lecturer in the Department of Zoology, University of Oxford, UK.

Multivariate biology

Robert R. Sokal

Morphometrics, The Multivariate Analysis of Biological Data. By Richard A. Pimentel. Pp.288. (Kendall/Hunt: Dubuque, Iowa, 1979.) Paperback \$19.95.

PRESENTING the results of morphometric, especially multivariate morphometric research, to a scientific audience is a major challenge. How can the reader make sense of numerous matrices, vectors and bivariate plots embellished with often unfamiliar terminology and esoteric mathematical and statistical manipulations? The author has set himself the task of remedying this situation by introducing biologists in a mathematically simple manner to the various multivariate techniques. To prevent misunderstandings I must point out at the outset that Pimentel uses the term morphometrics in an unconventional sense to include the application of multivariate analysis to unravel any pattern of variation including ecology, behaviour and physiology, not just the

biometric analysis of phenotypic morphological variation in animals and plants.

The standard topics of multivariate analysis are covered in a very personal style with the author providing caveats about use of various methods and indicating quite clearly which he prefers. Multiple factor analysis is given short shrift, whereas discriminant analysis is given a very full, well-presented treatment. Although there are some examples of applications, I would have liked to see a specific section in which the methods are compared and the types of methods suitable for given biological questions are pointed out. In my judgement, multivariate morphometrics (in the narrow sense) is one of the most promising disciplines of evolutionary research. Yet so far a general discussion of this subject from a biological point of view has not appeared. Thus in addition to this useful volume we need yet another account — a discussion of the genetic, developmental and ecological implications of the multivariate assessment of morphological phenotypes. □

Robert R. Sokal is Leading Professor of Ecology and Evolution at the State University of New York at Stony Brook.